Giant mimivirus R707 encodes a glycogenin paralogue polymerizing glucose through - and -glycosidic linkages

Rommel, Anna J ; Hülsmeier, Andreas J ; Jurt, Simon ; Hennet, Thierry

Abstract: Acanthamoeba polyphaga mimivirus is a giant virus encoding 1262 genes among which many were previously thought to be exclusive to cellular life. For example, mimivirus genes encode enzymes involved in the biosynthesis of nucleotide sugars and putative glycosyltransferases. We identified in mimivirus a glycogenin-1 homologous gene encoded by the open reading frame R707. The R707 protein was found to be active as a polymerizing glucosyltransferase enzyme. Like glycogenin-1, R707 activity was divalent-metal-ion-dependent and relied on an intact DXD motif. In contrast with glycogenin-1, R707 was, however, not self-glucosylating. Interestingly, the product of R707 catalysis featured 1-6, 1-6 and 1-4 glycosidic linkages. Mimivirus R707 is the first reported glycosyltransferase able to catalyse the formation of both and linkages. Mimivirus-encoded glycans play a role in the infection of host amoebae. Co-infection of Acanthamoeba with mimivirus and amylose and chitin hydrolysate reduced the number of infected amoebae, thus supporting the importance of polysaccharide chains in the uptake of mimivirus by amoebae. The identification of a glycosyltransferase capable of forming and linkages underlines the peculiarity of mimivirus and enforces the concept of a host-independent glycosylation machinery in mimivirus.

DOI: https://doi.org/10.1042/BCJ20160280

Giant mimivirus R707 encodes a glycogenin paralog polymerizing glucose through alpha and beta glycosidic linkages

Anna J. Rommel[1], Andreas J. Hülsmeier[1], Simon Jurt[2], and Thierry Hennet[1]*

[1]Institute of Physiology, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland; [2]Institute of Organic Chemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

*Corresponding author
E-mail: thierry.hennet@uzh.ch
Phone: +41 44 635 50 80

ABSTRACT

Acanthamoeba polyphaga mimivirus is a giant virus encoding 1'262 genes among which many were previously thought to be exclusive to cellular life. For example, mimivirus genes encode enzymes involved in the biosynthesis of nucleotide sugars and putative glycosyltransferases. We identified in mimivirus a glycogenin-1 homologous gene encoded by the open reading frame R707. The R707 protein was found to be active as a polymerizing glucosyltransferase enzyme. Like glycogenin-1 R707 activity was divalent metal ion dependent and relied on an intact DXD motif. In contrast to glycogenin-1, R707 was however not self-glucosylating. Interestingly, the product of R707 catalysis featured α1-6, β1-6, and α1-4 glycosidic linkages. Mimivirus R707 is the first reported glycosyltransferase able to catalyze the formation of both α and β linkages. Mimivirus encoded glycans play a role in the infection of host amoeba. Co-infection of *Acanthamoeba* with mimivirus and amylose and chitin hydrolysate reduced the number of infected amoeba, thus supporting the importance of polysaccharide chains in the uptake of mimivirus by amoeba. The identification of a glycosyltransferase capable of forming α and β linkages underlines the peculiarity of mimivirus and enforces the concept of a host-independent glycosylation machinery in mimivirus.

SUMMARY STATEMENT

We have identified the mimivirus R707 protein as paralog to glycogenin-1 with polymerizing glucosyltransferase activity, catalyzing the formation of α1-6, β1-6, and α1-4 glycosidic linkages. Mimivirus R707 is the first reported glycosyltransferase capable of forming both α and β linkages.

KEYWORDS: giant virus / polysaccharide / glycogen / glycosylation / nuclear magnetic resonance

SHORT TITLE: Glycogenin-1 paralog in giant mimivirus

ABBREVIATIONS

CID-MS	collision-induced dissociation mass spectrometry
DQF-COSY	double quantum filtered correlated spectroscopy
ESI-MS	electrospray ionization mass spectrometry
Gal	galactose
GC-MS	gas chromatography-mass spectrometry
Glc	glucose
GlcNAc	N-acetylglucosamine
GYG1	glycogenin-1
HCD-MS	higher-energy collisional dissociation mass spectrometry
HexNAc	N-acetylhexosamine
HMBC	heteronuclear multiple bind correlation
HSQC	heteronuclear single quantum coherence
MALDI-MS	matrix-assisted laser desorption/ionization mass spectrometry
NMR	nuclear magnetic resonance
PBCV-1	*Paramecium bursaria* chlorella virus
pNP	4-nitrophenyl
Rha	rhamnose
ROESY	rotating frame Overhauser enhancement spectroscopy
TOCSY	two-dimensional nuclear magnetic resonance spectroscopy
UDP	uridine diphosphate
Vio	viosamine
Xyl	xylose

INTRODUCTION

The giant *Acanthamoeba polyphaga* mimivirus is a member of the expanding family of *Mimiviridae*, belonging to the monophyletic group of nucleocytoplasmic large DNA viruses. The genome of mimivirus consisting of 1.2 Mbp and encoding 1'262 genes has been described in 2003 (1). The icosahedral mimivirus particle is 400 nm in diameter and covered with 140 nm long glycosylated protein fibers (2). Many mimivirus genes are unusual for viruses as they code for genes typical of cellular life, such as tRNA synthases, DNA repair factors, chaperones, and metabolic enzymes (3). Interestingly mimivirus encodes multiple genes possibly composing a functional glycosylation machinery.

Virus glycoproteins are associated with host-virion binding, virus entry, as for example hepatitis C virus E1/E2, and host-immune evasion as HIV gp120 (4). Glycosylation of virus envelope or surface proteins can promote proper folding, trafficking, and stability of virus proteins, such as for example in influenza virus hemagglutinin (4). In contrast to giant viruses conventional viruses hijack the glycosylation machinery of their host and, with few examples, do not encode glycosyltransferases or other genes involved in glycosylation (4). A giant virus encoding its own glycosylation machinery is found in the family of Chloroviruses. *Paramecium bursaria* chlorella virus (PBCV-1), a member of *Phycodnaviridae* that infects unicellular green algae, encodes at least five putative glycosyltransferases as well as multiple enzymes involved in sugar metabolism (5). The major capsid protein Vp54 of PBCV-1 carries a virus specific N-glycan that is added to the protein independently of the host's ER and Golgi pathway (5). This N-glycan is not attached to the canonical sequon and the glycan structure is atypical and more reminiscent of bacterial and archaeal glycans (6).

In mimivirus the biosynthetic pathways of three nucleotide sugars are well characterized. Mimivirus synthesizes UDP-viosamine (Vio), UDP-rhamnose (Rha), and UDP-N-acetylglucosamine (GlcNAc) independently of its host *Acanthamoeba polyphaga* (7-9). These nucleotide sugar donors reflect the monosaccharide composition of mimivirus, given that GlcNAc, Rha, and Vio are major carbohydrates isolated from mimivirus particles in addition to glucose (Glc) (8, 10). The analysis of the O-glycan content of mimivirus shows various glycan structures demanding for glycosyltransferases with different functionalities. Mimivirus proteins are glycosylated with linear O-glycans consisting of hexose polymers of variable length. Also branched O-glycans have been identified, which consist of a hexose at the reducing end, a terminal N-acetylhexosamine (HexNAc) residue on one branch and a terminal pentose residue linked to HexNAc on the second branch (10).

Analysis of the mimivirus genome suggests that in addition to the nucleotide-diphospho sugar synthesizing machinery mimivirus encodes at least 11 putative glycosyltransferases (3). These putative glycosyltransferases share sequence similarities with archaeal, bacterial, and eukaryotic glycosyltransferases. For example, the mimivirus L230 gene encodes a bi-functional collagen-modifying

enzyme structurally similar to PLOD3 lysyl hydroxylase and GLT25D1 galactosyltransferase of animal origin (11). Another mimivirus protein, R707, is structurally related to glycogenin-1 (GYG1), which initiates glycogen biosynthesis in animals. Despite this similarity, R707 unlikely mediates glycogen formation in mimivirus, but may be involved in the assembly of polysaccharide decorating the virus capsid. The present study characterizes the glycosyltransferase activity of mimivirus R707 and the importance of polysaccharides in the infection of amoeba by mimivirus.

EXPERIMENTAL/ MATERIALS AND METHODS

Expression and purification of recombinant R707 protein – Recombinant R707 was expressed in *Escherichia coli* BL21 De3 (Novagen, Nottingham, UK) as His_{10}-fusion protein in pET16b (Merck, Billerica, MA). Viral DNA was purified using phenol/chloroform (Roth, Karlsruhe, Germany). The DNA sequence corresponding to mimivirus R707 (RefSeq ID NC_014649.1) was amplified using the primers 5'-ATACTCGAGTCTTCCTATGCATATGTTACAG-3' and 5'-ATAGGATCCTTAATAAGGTAG-TTTAATGTCA-3' containing XhoI and BamHI restriction sites, respectively (underlined bases). Site-directed mutagenesis of R707 was performed by QuickChange site-directed mutagenesis (12) using primers listed in table 1. Asn position 103 was substituted for Ser, Tyr 215 was substituted for Phe, and Lys 232 was substituted for Arg. For protein expression 100 ml lysogeny broth containing 100 µg/ml ampicillin (Sigma-Aldrich, Buchs, Switzerland) were inoculated with 1 ml of overnight culture. Protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (Biosolve, Dieuze, France) when bacteria reached OD_{600} value of 0.4. *E. coli* were grown at 16°C for 16 h, shaking, then pelleted at 6'000 x g, resuspended in 15 ml 500 mM NaCl, 20 mM Tris pH 7.4, 10% v/v glycine, 10 mM imidazole (MCAC10) and lysed using a French Press (Avestin, Mannheim, Germany). His_{10}-tagged R707 was purified from the soluble fraction using 100 µl Ni-Sepharose 6 Fast Flow beads (GE Healthcare, Glattbrugg, Switzerland). After binding to beads in MCAC10 at 4°C for 4 h, rotating, R707 was eluted in 400 mM imidazole (MCAC400). Purified R707 was stored at 4°C.

Enzymatic activity – Glycosyltransferase activity was measured in 20 mM Tris-HCl pH 7.4 containing 0.5 µl 20 µCi/ml UDP-[^{14}C]-sugar (Perkin-Elmer, Waltham, MA), 20 µM UDP-sugar, 10 mM $MnCl_2$ (Sigma-Aldrich), 1 mM DTT, 50 mM 4-nitrophenyl (pNP)-sugar in DMSO (Sigma-Aldrich) as acceptor substrate, and 10 µg of purified R707 protein. The UDP-sugars tested as donor substrates were UDP-Glc, UDP-GlcNAc, UDP-galactose (Gal), and UDP-xylose (Xyl). The pNP-sugars tested as acceptor substrates were pNP-Glc, pNP-Xyl, pNP-Gal, and pNP-Rha (Sigma-Aldrich). Assays were incubated at 35°C for 4 h, then stopped by addition of ice cold 500 µl H_2O, and purified over SepPak-C18 cartridges (Waters, Milford, MA). Samples were added to the cartridges, washed with 10 ml of H_2O, and eluted with 5 ml of methanol

(Sigma-Aldrich). Radioactivity was measured after addition of 10 ml of IRGASAFE scintillation liquid (Perkin-Elmer) in a Tri-Carb 2900TR scintillation counter (Perkin-Elmer). Self-glucosylation of R707 was measured as described above, but without addition of pNP-sugar acceptor substrate. After incubation assays were stopped by addition of 500 µl 5% trichloroacetic acid / 5% phosphotungstic acid and incubated on ice for 30 min. Precipitates were recovered on 1 µm glass filters (VWR, Radnor, PA) using a vacuum manifold and radioactivity was measured as described above.

Product analysis of R707 glycosyltransferase assay – About 200 pmol of glycosyltransferase assay product were subjected to MALDI-MS analysis. MALDI-MS were acquired essentially as described earlier (13). For NMR analysis assay products from 250 glycosyltransferase assays were pooled, dried, and resuspended in 300 µl DMSO. To remove spare pNP-α-Glc, products were purified by gel-filtration using a Superdex Peptide HR 10/30 (GE Healthcare) and 10% acetonitrile, 0.1% trifluoroacetic acid at a flow rate of 0.1 ml/min. Different product species were separated using a Hypersil C18 column (Thermo Scientific, Waltham, MA) under the following conditions: buffer A: 0.1% trifluoroacetic acid (Sigma-Aldrich), buffer B: 0.1% trifluoroacetic acid, 10% acetonitrile (Sigma-Aldrich), 30 min 97.5% buffer A/ 2.5% buffer B, 120 min gradient to 10% buffer B, flow rate: 0.5 ml/min. Product elution was monitored at 375 nm, and 2 min-fractions were collected. Peaks containing a sufficient amount of product were pooled, dried, and subjected to NMR analysis. Products were dissolved in 220 µl D_2O and transferred into 5 mm D_2O-matched Shigemi NMR tubes (Shigemi, Allison Park, PA). Structure elucidation of sample A and B is based on a set of 1D proton NMR, 2D ^{1}H-^{1}H double quantum filtered correlated spectroscopy (DQF-COSY), ^{1}H-^{1}H two-dimensional nuclear magnetic resonance spectroscopy (TOCSY) (100 ms mixing time), ^{1}H-^{1}H rotating frame Overhauser enhancement spectroscopy (ROESY) (300 ms mixing-time), ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) and ^{1}H-^{13}C heteronuclear multiple bind correlation (HMBC) experiments. Identification of sample C is based on 1D ^{1}H and 2D ^{1}H-^{13}C HSQC spectra that were also recorded for the reference compounds, namely pNP-α-maltose (Sigma-Aldrich). All spectra were collected at 298 K on a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenically cooled 5 mm TCI probe and externally referenced against 4,4-dimethyl-4-silapentane-1-sulfonate with a $^{13}C/^{1}H$ ratio of 0.25144953 used for the indirect ^{13}C referencing (14). Spectra were processed using the Bruker TopSpin 3.5 software suite. For linkage analysis of the products A, B, and C GC-MS experiments were done essentially as described previously (10). The initial oven temperature of 80°C was held for 2 min, increased by 8°C/min until 320°C and held for 8 min. In addition, the products A, B, and C were analyzed by negative-ion ESI-MS and ESI-HCD-MS/MS experiments were carried out on an Orbitrap Fusion mass spectrometer (Thermo Scientific), analogous to the negative ion ESI-CID-MS/MS method described recently (15). The glycosyltransferase reaction products were dissolved in 50 % acetonitrile, 2 mM NH_4HCO_3 (Sigma-Aldrich) at a concentration of 30 µM and 2 µl were injected to the mass spectrometer. Solvent 50 % acetonitrile, 2 mM NH_4HCO_3 was delivered by a Harvard syringe pump at a flow rate of 4 ml/min. The ion- transfer tube temperature was

maintained at 320°C and the capillary voltage was maintained at – 2700 V. For MS/MS acquisition quadrupole isolation with a 2 m/z isolation width and HCD activation with 0 % collisional energy was used. Ions were detected in the Orbitrap at a resolution of 15'000 and a scan range of 50-500 m/z. The AGC target was set to 5 x 10^4 and maximal injection time to 100 ms.

Phagocytosis competition assay – Amoeba (5 x 10^3) were seeded in 96-well plates in peptone yeast glucose broth and infected at a multiplicity of 250 with purified mimivirus particles together with 1 to 10 µg of amylose, chitin hydrolysate, dextran, or mannan (Sigma-Aldrich). Polyvinyl alcohol (Sigma-Aldrich) was used as negative control. Amoeba were incubated with mimivirus and polymers at 28°C for 4.5 h, and then fixed in 4% paraformaldehyde. Intracellular DNA including virus factories was stained with DAPI for 30 min. Virus factories per total number of amoeba per view field were quantified using an Axiovert 200M microscope (Zeiss, Oberkochen, Germany) and Fiji cell counter (16).

Statistical analysis - ANOVA with multiple comparisons was used and a P-value of $P < 0.05$ was considered as statistically significant. GraphPad Prism 6 (GraphPad Software, Inc., La Jolla, CA) was used for statistical analysis. Results are presented as mean ± SEM.

RESULTS

The mimivirus R707 protein was identified as putative glycosyltransferase based on sequence similarity to animal GYG1, which primes glycogen synthesis by initiated self-glucosylation (17). GYG1 transfers Glc from UDP-Glc to a conserved Tyr residue at position 195, then elongates this Glc residue with up to seven additional Glc units through α1,4-linkage, making GYG1 a retaining enzyme (17). R707 and GYG1 shared 26% sequence identity and 42% similarity. The mimivirus R707 protein included a DXD motif, which aligns with the same motif in animal GYG1 proteins (Figure 1A). The DXD motif is typically found in glycosyltransferases utilizing nucleotide-activated sugars as donor substrates, and is involved in the coordination of a divalent metal ion in the catalytic pocket of the enzyme (18). In GYG1 Leu position 8, Thr 10, Asn 11, Tyr 14, Asp 101, Ala 102, Asp 103, Asn 132, Gln 163, His 211, Gly 214, and Lys 217 are involved in the coordination of UDP-Glc. The homologous residues Asn 13, Tyr 16, Asp 101, Asp 103, Asn 150, His 226, and Lys 232 were also found to be conserved in mimivirus R707 (Figure 1B) (19). The pattern of sequence and motif similarity between GYG1 and R707 suggested a common Glc polymerizing activity for both proteins. Based on BLAST analysis the putative glycosyltransferase domain spanned nearly the entire R707 protein between amino acid position 4 and 265 (Figure 1C). A 3D-model comparison of R707 and GYG1 using PHYRE[2] and SWISS-MODEL shows strong conservation around the DXD motif and main secondary structures. The majority of the α-helices and β-sheets of GYG1 from residue 1-230 align with R707 (20, 21). Both proteins show structural diversity towards the C-terminus as predicted by BLAST

analysis. The R707 gene is expressed in the intermediate and late phase of mimivirus replication, indicating a possible involvement of R707 in virion maturation and fiber formation (22).

The R707 protein was expressed in *E. coli* as His_{10}-fusion protein, purified by affinity chromatography, and glycosyltransferase activity was tested. For product purification pNP-α-Glc and pNP-β-Glc were used as acceptor substrates. In presence of pNP-α-Glc R707 showed a specific activity of 19.6 pmol/min/mg. In presence of pNP-β-Glc R707 activity reached only 7.1 pmol/min/mg (Figure 2A). To test for a possible self-glucosylation activity of R707, we omitted the acceptor substrate from the enzymatic reaction. We could not detect any self-glucosylation activity on R707 (Figure 2A).

To assess the functionality of different residues of R707, we mutated amino acids in R707, which correspond to the essential residues Asp 103, Tyr 195, and Lys 217 in GYG1. The residue Asp 103 of R707 is part of the conserved DXD motif, Tyr 215 corresponds to the glucosylated Tyr residue of GYG1, and Lys 232 corresponds to a residue involved in the coordination of UDP-Glc in GYG1 (19). The loss of the DXD motif and of the coordinating Lys 232 residue abolished the glycosyltransferase activity of R707. By contrast, the mutation of Tyr 215 had no effect on R707 activity (Figure 2B), confirming the lack of self-glycosylation in the mimivirus glycosyltransferase. Like GYG1, the glycosyltransferase activity of R707 required the divalent cation Mn^{2+}. R707 activity was decreased to 19% in absence of Mn^{2+}, to 20% after substitution of Mn^{2+} with Mg^{2+}, and to 10% after substitution of Mn^{2+} with Ca^{2+} (Supplementary Figure 1A). In GYG1 the substitution of Mn^{2+} with Mg^{2+} and Ca^{2+} decreased glycosyltransferase activity to 5% and 10%, respectively (19).

To determine the substrate specificity of R707, we tested the nucleotide-sugar donors UDP-Glc, UDP-GlcNAc because these monosaccharides are the main carbohydrates of mimivirus glycans (8, 10). We also tested UDP-Gal as an epimer of UDP-Glc and UDP-Xyl because it is used as alternative nucleotide-sugar donor by GYG1 (17). Similarly, we tested the monosaccharides pNP-α-Glc, pNP-α-Xyl, pNP-α-Gal, and pNP-α-Rha as acceptor substrates. UDP-Glc was the only donor substrate compatible with R707 glucosyltransferase activity (Supplementary Figure 1B). The corresponding glucosyltransferase activity was highest towards the α-linked Glc acceptor pNP-α-Glc, but R707 showed also significant glucosyltransferase activity towards pNP-α-Xyl. By contrast, none of the other monosaccharides tested were valid acceptor substrates for R707 (Supplementary Figure1 C). As a comparison, GYG1 can also use UDP-Xyl as donor substrate, although the transfer of Xyl blocks the elongation of the Glc polymer by GYG1 (17).

To characterize the product of R707 catalysis we first subjected 200 pmol of purified reaction product to mass spectrometry analysis. In addition to the acceptor substrate pNP-α-Glc, the elongation products pNP-α-Glc_2 and pNP-α-Glc_3 were detected with a ppm error < 50 ppm (Figure 3), confirming the polymerizing activity of R707. A fragment spectrum of pNP-α-Glc_2 confirmed the presence of the product. To analyze the type of glycosidic linkage catalyzed by R707, reaction products were purified by reverse-phase

chromatography allowing a qualitative separation of the main products of the glycosyltransferase reaction. Two main chromatography peaks plus one additional peak recovered were subjected to NMR analysis (Figure 4). All other analyzed peaks showed a too high degree of structural heterogeneity to enable NMR analysis and linkage determination. The glycosidic linkages were identified from the cross peaks in the ROESY spectra. The rotating-frame Overhauser enhancements between the anomeric proton 1' and the protons 6a/6b established the 1,6-linkage for the two main reaction products analyzed (Figure 5, Supplementary Figure 2). Differentiation between the α- and β-configuration at the anomeric centers was based on the $^{3}J_{1,2}$ coupling constants. Except for the anomeric proton 1' of sample A these couplings could be directly determined from the splitting of the H1 (H1') doublets in the 1D proton spectrum. For the anomeric proton 1' of sample A, which was obscured by the HOD signal, this coupling was identified from the H2'/H1' cross peak patterns in the DQF-COSY and then extracted from the H2' multiplet in the 1D proton spectrum. For comparison all the H2/H1 (H2'/H1') COSY cross peaks are shown (Figure 5). The neighboring coupling constant between anomeric protons indicates their relative orientation in pyranose structures. In a β configuration a coupling constant of 7-8 Hz is observed, in an α configuration the coupling constant is around 4 Hz (23). The following coupling constants were obtained: sample A: $^{3}J_{1,2}$=3.7 Hz; $^{3}J_{1',2'}$ =3.7 Hz; sample B: $^{3}J_{1,2}$=3.6 Hz; $^{3}J_{1',2'}$ =8.0 Hz. The ^{3}J coupling constants confirmed the α-configuration for product A and the β-configuration for product B (Figure 5). The additional reaction product C of R707 was also subjected to NMR analysis; a comparison of the HSQC data with the authentic standard indicated its identity as pNP-α-Glc(α1-4)Glc (Supplementary Figure 3). The NMR analysis thus identified pNP-α-Glc(α1-6)Glc, pNP-α-Glc(β1-6)Glc, and pNP-α-Glc(α1-4)Glc as R707 reaction products. Peaks A, B, and C were further analyzed by GC-MS and negative ion ESI-MS/MS for linkage determination of the pNP-disaccharides. The elution positions in the GC-MS analysis and the fragment ion products of the negative-ion ESI-MS identified a Glc(1-6)Glc linkage for peaks A and B and Glc(1-4)Glc linkage for peak C, corroborating the assignments based on NMR analyses (Figure 6, Supplementary Figure 4).

Mimivirus glycans are mainly found on the heavily glycosylated surface fibers (2, 8). Any additional role for mimivirus glycans in a later stage of infection, however, cannot be ruled out. Considering the formation of linear Glc polymers bearing various linkages by R707, we investigated the functional significance of such Glc polymers and other polysaccharides in the context of mimivirus infection. Upon infection of amoeba by mimivirus, we detected virus factories in 77% of amoeba 4.5 h post infection (Figure 7A). The addition of the α1-4 linked Glc polymer amylose to amoeba decreased mimivirus infection in a dose dependent manner, indicating that α1-4 linked Glc chains are indeed required for the uptake of the virus by amoeba (Figure 7B). A similar inhibition of virus infection was observed by addition of β1-4 linked GlcNAc in the form of chitin hydrolysate. The localization of the major proportion of GlcNAc is unknown to date as only a small proportion of O-glycan structures contain GlcNAc even though GlcNAc is the major

monosaccharide in mimivirus (10). By contrast, addition of the β1-6 linked Glc polymer dextran and of the α1-4 linked mannose polymer mannan had no effect on mimivirus infection (Figure 7B). The competitive inhibition of mimivirus infection by amylose and chitin hydrolysate confirmed the importance of polysaccharides in the capture of mimivirus by amoeba.

DISCUSSION

We have shown that mimivirus R707 is a functional glycosyltransferase transferring Glc from UDP-Glc to Glc through multiple glycosidic linkages. The R707 enzyme has strict donor and acceptor substrate specificities and its glucosyltransferase activity depends on Mn^{2+}. The characterized products of R707 reaction were pNP-α-Glc(α1-6)Glc, pNP-α-Glc(β1-6)Glc, and pNP-α-Glc(α1-4)Glc. Infection of *Acanthamoeba* with mimivirus in presence of soluble amylose and chitin hydrolysate reduced the number of infected amoeba indicating a role for glycan-binding protein interactions in the infection of amoeba.

This is the first report of a single-domain glycosyltransferase capable of forming both α- and β-glycosidic linkages. With few exceptions glycosyltransferases exhibit strict linkage specificities. In some unique cases glycosyltransferases exhibit flexibility regarding the formed glycosidic bond. Human fucosyltransferase III attaches fucose in either α1-3 or α1-4 linkage (24). The α-dystroglycan modifying enzyme LARGE is a bi-functional glycosyltransferase with both xylosyltransferase and glucuronyltransferase activities. LARGE generates repeats of the disaccharide unit (β1-3)Xyl(α1-3)glucuronic acid. In contrast to R707, LARGE consists of two separate glycosyltransferase domains with each domain responsible for a specific activity (25). We found 58% identity in amino acid residues in the catalytic pocket of GYG1 and R707. The remaining different residues might play a role in directing the inverting and retaining mechanism of linkage formation. In inverting glycosyltransferases an active-site side chain is needed as base catalyst for deprotonation the incoming nucleophile of the acceptor substrate. In retaining glycosyltransferases a nucleophile within the active site is needed to allow a double-displacement reaction (26). Crystallization of R707 and further studies of the catalytic pocket in presence and absence of the acceptor substrate may identify the residues possibly serving as base catalyst and nucleophilic side chains in the transfer of Glc. This could explain the flexibility of R707 in mediating distinct linkages and configurations.

GYG1 catalysis produces linear α1-4 Glc polymers. Branching occurs every 8-10 residues and is catalyzed by the glycogen branching enzyme GBE1, which catalyzes the α1-6 addition of Glc (27). We identified pNP-α-Glc(α1-6)Glc and pNP-α-Glc(β1-6)Glc and pNP-α-Glc(α1-4)Glc as products of R707 catalysis. The Glc trimers we identified were linear chains of Glc. This goes in hand with data published on mimivirus O-glycans, where all identified hexose polymers were linear (10). Glycogen is the universal molecule for Glc storage. Glycogen synthesis has been reported for a large number of species, from bacteria to yeast to mammals (17, 28, 29). In mammals, glycogen synthesis is initiated by self-glucosylation of GYG1 in the

cytosol (27, 29). In contrast to GYG1 mimivirus R707 did not exhibit self-glucosylating activities and a disruption of the Tyr215 residue, corresponding to Tyr 195 in GYG1, had no influence on R707 activity. Mimivirus depends on energy supply from amoeba and as the virus needs immediate energy availability energy storage in form of glycogen hardly makes sense (30). It is more likely that GYG1 and R707 showed divergent evolution where GYG1 evolved to being the core protein of glycogen and R707 to being a glycosyltransferase involved in decorating other mimivirus proteins with Glc polymers.

An independent glycosylation machinery allows mimivirus to form virus-specific glycan structures. These specific glycans are probably needed for virus-host interactions and mimivirus infection. Mimivirus polysaccharides could mimic the cell wall of gram-positive bacteria and thereby increase the chance of phagocytosis by *Acanthamoeba* (31-33). *Acanthamoeba spp.* express various lectins of which some are known to be involved in phagocytosis of their prey (34-36). Parallel addition of amylose or chitin hydrolysate during mimivirus infection of amoeba reduces infection in a dose dependent manner. This effect is not seen in presence of dextran, mannan or poly vinyl alcohol suggesting specific receptor occupancy on the surface of amoeba or mimivirus resulting in reduced mimivirus uptake. In addition, flexible and diverse surface glycosylation might allow mimivirus to target different host species, as observed for *Legionella pneumophila*. *Legionella pneumophila* is taken up by both *Acanthamoeba castellanii* and *Naegleria lovaniensis* via receptor-mediated endocytosis but through binding to the different receptors, mannose-binding receptor on the one hand, and N-acetyl-D-galactosamine receptor on the other hand (36).

The results described in this study present a basis for future investigations for a better understanding of the mechanism of glycosidic bond formation in glycosyltransferases. Our results further establish the existence of a host-independent glycosylation machinery in mimivirus. Furthermore, the findings presented underline the importance of glycan-protein interactions in mimivirus infection.

ACKNOWLEDGMENTS

We thank Luca Plan for his assistance with the mutagenesis experiments.

DECLARATION OF INTEREST

The authors declare no conflict of interest with the content of this article.

FUNDING INFORMATION

This work was supported by the Research Credit of the University of Zurich to AR and by the Swiss National Foundation grant 310030_149949 to TH.

AUTHOR CONTRIBUTION STATEMENT

Anna Rommel, Andreas Hüslmeier, and Thierry Hennet designed the experiments. Anna Rommel and Thierry Hennet wrote the manuscript with contributions from Andreas Hülsmeier and Simon Jurt. Andreas Hüslmeier carried out all MS analyses. Simon Jurt did the NMR experiments. All other experiments were done by Anna Rommel.

REFERENCES

1. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, et al. A giant virus in amoebae. Science. 2003;299(5615):2033. Epub 2003/03/29.
2. Kuznetsov YG, Xiao C, Sun S, Raoult D, Rossmann M, McPherson A. Atomic force microscopy investigation of the giant mimivirus. Virology. 2010;404(1):127-37. Epub 2010/06/17.
3. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of Mimivirus. Science. 2004;306(5700):1344-50. Epub 2004/10/16.
4. Vigerust DJ, Shepherd VL. Virus glycosylation: role in virulence and immune interactions. Trends in microbiology. 2007;15(5):211-8. Epub 2007/04/03.
5. Van Etten JL, Gurnon JR, Yanai-Balser GM, Dunigan DD, Graves MV. Chlorella viruses encode most, if not all, of the machinery to glycosylate their glycoproteins independent of the endoplasmic reticulum and Golgi. Biochimica et biophysica acta. 2010;1800(2):152-9. Epub 2009/08/06.
6. De Castro C, Molinaro A, Piacente F, Gurnon JR, Sturiale L, Palmigiano A, et al. Structure of N-linked oligosaccharides attached to chlorovirus PBCV-1 major capsid protein reveals unusual class of complex N-glycans. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(34):13956-60. Epub 2013/08/07.
7. Parakkottil Chothi M, Duncan GA, Armirotti A, Abergel C, Gurnon JR, Van Etten JL, et al. Identification of an L-rhamnose synthetic pathway in two nucleocytoplasmic large DNA viruses. Journal of virology. 2010;84(17):8829-38. Epub 2010/06/12.
8. Piacente F, Marin M, Molinaro A, De Castro C, Seltzer V, Salis A, et al. The giant DNA virus Mimivirus encodes a pathway for the biosynthesis of the unusual sugar 4-amino-4,6-dideoxy-D-glucose (viosamine). The Journal of biological chemistry. 2011. Epub 2011/12/14.
9. Piacente F, Bernardi C, Marin M, Blanc G, Abergel C, Tonetti MG. Characterization of a UDP-N-acetylglucosamine biosynthetic pathway encoded by the giant DNA virus Mimivirus. Glycobiology. 2014;24(1):51-61. Epub 2013/10/11.
10. Hulsmeier AJ, Hennet T. O-Linked glycosylation in Acanthamoeba polyphaga mimivirus. Glycobiology. 2014;24(8):703-14. Epub 2014/05/06.
11. Luther KB, Hulsmeier AJ, Schegg B, Deuber SA, Raoult D, Hennet T. Mimivirus collagen is modified by a bifunctional lysyl hydroxylase and glycosyltransferase enzyme. The Journal of biological chemistry. 2011. Epub 2011/11/03.
12. Braman J, Papworth C, Greener A. Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol Biol. 1996;57:31-44. Epub 1996/01/01.
13. Hulsmeier AJ, Welti M, Hennet T. Glycoprotein maturation and the UPR. Methods in enzymology. 2011;491:163-82. Epub 2011/02/19.
14. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, et al. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. Journal of biomolecular NMR. 1995;6(2):135-40. Epub 1995/09/01.
15. Palma AS, Liu Y, Zhang H, Zhang Y, McCleary BV, Yu G, et al. Unravelling glucan recognition systems by glycome microarrays using the designer approach and mass spectrometry. Mol Cell Proteomics. 2015;14(4):974-88.
16. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature methods. 2012;9(7):676-82. Epub 2012/06/30.
17. Lomako J, Lomako WM, Whelan WJ. Glycogenin: the primer for mammalian and yeast glycogen synthesis. Biochimica et biophysica acta. 2004;1673(1-2):45-55. Epub 2004/07/09.
18. Ramakrishnan B, Boeggeman E, Ramasamy V, Qasba PK. Structure and catalytic cycle of beta-1,4-galactosyltransferase. Current opinion in structural biology. 2004;14(5):593-600. Epub 2004/10/07.
19. Gibbons BJ, Roach PJ, Hurley TD. Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. Journal of molecular biology. 2002;319(2):463-77. Epub 2002/06/08.
20. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nature protocols. 2009;4(1):1-13. Epub 2009/01/10.

21. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols. 2015;10(6):845-58. Epub 2015/05/08.
22. Legendre M, Santini S, Rico A, Abergel C, Claverie JM. Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. Virology journal. 2011;8:99. Epub 2011/03/08.
23. Duus J, Gotfredsen CH, Bock K. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. Chem Rev. 2000;100(12):4589-614. Epub 2001/12/26.
24. Rini J, Esko J, Varki A. Glycosyltransferases and Glycan-processing Enzymes. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, et al., editors. Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY)2009.
25. Inamori K, Yoshida-Moriguchi T, Hara Y, Anderson ME, Yu L, Campbell KP. Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. Science. 2012;335(6064):93-6. Epub 2012/01/10.
26. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. Annual review of biochemistry. 2008;77:521-55. Epub 2008/06/04.
27. Smythe C, Cohen P. The discovery of glycogenin and the priming mechanism for glycogen biogenesis. European journal of biochemistry / FEBS. 1991;200(3):625-31. Epub 1991/09/15.
28. Okita TW, Rodriguez RL, Preiss J. Biosynthesis of bacterial glycogen. Cloning of the glycogen biosynthetic enzyme structural genes of Escherichia coli. The Journal of biological chemistry. 1981;256(13):6944-52. Epub 1981/07/10.
29. Wilson WA, Roach PJ, Montero M, Baroja-Fernandez E, Munoz FJ, Eydallin G, et al. Regulation of glycogen metabolism in yeast and bacteria. FEMS microbiology reviews. 2010;34(6):952-85. Epub 2010/04/24.
30. Mutsafi Y, Fridmann-Sirkis Y, Milrot E, Hevroni L, Minsky A. Infection cycles of large DNA viruses: Emerging themes and underlying questions. Virology. 2014. Epub 2014/07/06.
31. Klose T, Herbst DA, Zhu H, Max JP, Kenttamaa HI, Rossmann MG. A Mimivirus Enzyme that Participates in Viral Entry. Structure. 2015;23(6):1058-65. Epub 2015/05/20.
32. Sobhy H, Scola BL, Pagnier I, Raoult D, Colson P. Identification of giant Mimivirus protein functions using RNA interference. Frontiers in microbiology. 2015;6:345. Epub 2015/05/15.
33. Rodrigues RA, Dos Santos Silva LK, Dornas FP, de Oliveira DB, Magalhaes TF, Santos DA, et al. Mimivirus Fibrils Are Important for Viral Attachment to the Microbial World by a Diverse Glycoside Interaction Repertoire. Journal of virology. 2015;89(23):11812-9. Epub 2015/09/18.
34. Clarke M, Lohan AJ, Liu B, Lagkouvardos I, Roy S, Zafar N, et al. Genome of Acanthamoeba castellanii highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling. Genome biology. 2013;14(2):R11. Epub 2013/02/05.
35. Medina G, Flores-Martin S, Fonseca B, Otth C, Fernandez H. Mechanisms associated with phagocytosis of Arcobacter butzleri by Acanthamoeba castellanii. Parasitology research. 2014;113(5):1933-42. Epub 2014/03/22.
36. Declerck P, Behets J, De Keersmaecker B, Ollevier F. Receptor-mediated uptake of Legionella pneumophila by Acanthamoeba castellanii and Naegleria lovaniensis. Journal of applied microbiology. 2007;103(6):2697-703. Epub 2007/09/14.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of molecular biology. 1990;215(3):403-10. Epub 1990/10/05.
38. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic acids research. 2015;43(Database issue):D222-6. Epub 2014/11/22.
39. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. 2011;7:539. Epub 2011/10/13.

FIGURE LEGENDS

FIGURE 1. Sequence homology between mimivirus R707 and eukaryotic glycogenin-1 proteins. (A) Alignment of R707 and glycogenin-1 (GYG1) from different species using protein BLAST and the non-redundant protein sequences database (37, 38). The multiple sequence alignment of R707 and GYG1 homologs was generated using Clustal Omega (39). Identical residues are shaded in black (≥ 4 species) or in grey (<4 species). The DXD motif and catalytic tyrosine of GYG1 are marked with dashed boxes. Mutated residues of R707 are marked with open triangles. (B) Catalytic center of mouse Gyg1 coordinating UDP-Glc and Mn^{2+} (adapted from (19)). Residues conserved in R707 are marked in blue. (C) Schematic representation of full-length R707 (281 amino acids long) and of its predicted glycosyltransferase domain spanning residues 4 to 265.

FIGURE 2. Glycosyltransferase activity of R707. (A) Glucosyltransferase activity of recombinant R707 on pNP-α-Glc, pNP-β-Glc and self-glucosylation of R707 in absence of an acceptor substrate. (B) Glucosyltransferase activity of wildtype R707 in comparison to R707 including the amino acid substitutions D103S, Y215F, and K232R. Two independent experiments with triplicates in each instance were done. Results are shown as mean ± SEM.

FIGURE 3. (A) MALDI-MS analysis of R707 assay products. 200 pmol of glycosyltransferase assay product were subjected to MALDI-MS analysis. The molecular ions of assay products are labeled with their glycan annotation. Glucose residues are represented by open circle. Ppm error (pNP-Glc_2) = 34 ppm, ppm error (pNP-Glc_3) = 15 ppm. (B) MS2 spectrum of m/z 486.1. The inserted symbols represent the proposed assignments.

FIGURE 4. R707 product purification. Approximately 300 nmol of R707 assay products were separated on a Hypersil C18 column. Peaks A, B, and C were subjected to further structural analysis by NMR.

FIGURE 5. Linkage analysis of R707 assay product by NMR with (A) referring to peak A, figure 4 and (B) referring to peak B from figure 4. Two unidentified product species were subjected to proton NMR TOCSY (middle panel, top) for identification of the two spin systems from the sugar units. Configuration between pNP and Glc (COSY, lower panel) and configuration between Glc and Glc (COSY, upper panel) by expansion of the DQF-COSY cross peaks between the anomeric proton and proton at position 2 or 2' respectively allowing extraction of the ^3J coupling constants and discrimination between the alpha ($^3J_{(1,2)}$~3 Hz) and beta ($^3J_{(1,2)}$~7 Hz) configuration. Rotating-frame Overhauser enhancement between aromatic proton A and anomeric proton 1 (ROESY, left panel) for linkage determination between pNP and Glc. Rotating-frame Overhauser enhancement between anomeric proton 1' and protons 6a/6b (ROESY, middle panel, bottom) for linkage determination between Glc and Glc. The experiment was done twice.

FIGURE 6. Product linkage analysis of peaks A, B, and C using GC-MS. Partially methylated alditol acetates (PMMA) were prepared from the R707 reaction products and resolved by GC-MS. Extracted chromatograms at m/z 118.061. Characteristic hexoses are shown. The elution positons of 2-, 3-, 4-, and 6-O-substituted PMMA were determined by comparison with PMMAs prepared from authentic reference compounds.

FIGURE 7. Mimivirus uptake competition assay. (A) Amoeba were infected at a multiplicity of 250, (B) and in combination with 1 µg, 5 µg, or 10 µg of polysaccharides or poly vinyl alcohol (PVA). Virus factories per total number of amoeba per viewfield were quantified 4.5 h post infection using an Axiovert 200M microscope and Fiji cell counter. Two independent experiments quantifying five viewfields per condition were done. Results are shown as mean ± SEM.

SUPPLEMENTARY FIGURE 1. Co-factor and substrate specificity of R707. (A) Assays with Mn^{2+}, Mg^{2+}, or Ca^{2+} and in absence of divalent metal ions. (B) Nucleotide sugar donor specificity using UDP-Glc or UDP-GlcNAc or UDP-Gal or UDP-Xyl. (C) Acceptor substrate specificity using pNP-α-Glc, pNP-α-Xyl, pNP-α-Gal, or pNP-α-Rha. Two independent experiments with triplicates in each instance were performed. Results are shown as mean ± SEM.

SUPPLEMENTARY FIGURE 2 Heteronuclear single quantum coherence (HSQC) signals of peaks A and B

SUPPLEMENTRAY FIGURE 3. Heteronuclear single quantum coherence (HSQC) signals of peak C and pNP-maltose.

SUPPLEMENTARY FIGURE 4. Product analysis of peaks A, B, and C by negative-ion ESI-MS. 60 pmol of products were subjected to analysis. The inserted symbols represent the proposed assignments.

TABLE I PCR primers used to construct mutated R707. Underlined bases mark the introduced mutations. F: forward primer, R: reverse primer.

Mutated Residue	Primer sequence
mR707-Asp103Ser	F 5'-CGATAAAATTATTTTATTAGATTTAAGCATGATAATTGC-3'
	R 5'-GCAATTATCATGCTTAAATCTAATAAAATAATTTTATCG-3'
mR707-Tyr215Phe	F 5'-GGGTTGACACATCGAGTAAAAAAGTTCCATTATACAATTGA-3'
	R 5'-CAATTGTATAATGGAACTTTTTTACTCGATGTGTCAACCC-3'
mR707-Lys232Arg	F 5'-CATTTTTCTAGTTCCTATCGACCATGGAACAGATTAAATTCCGAC-3'
	R 5'-CGGAATTTAATCTGTTCCATGGTCGATAGGAACTAGAAAAATGA-3'

Figure 1

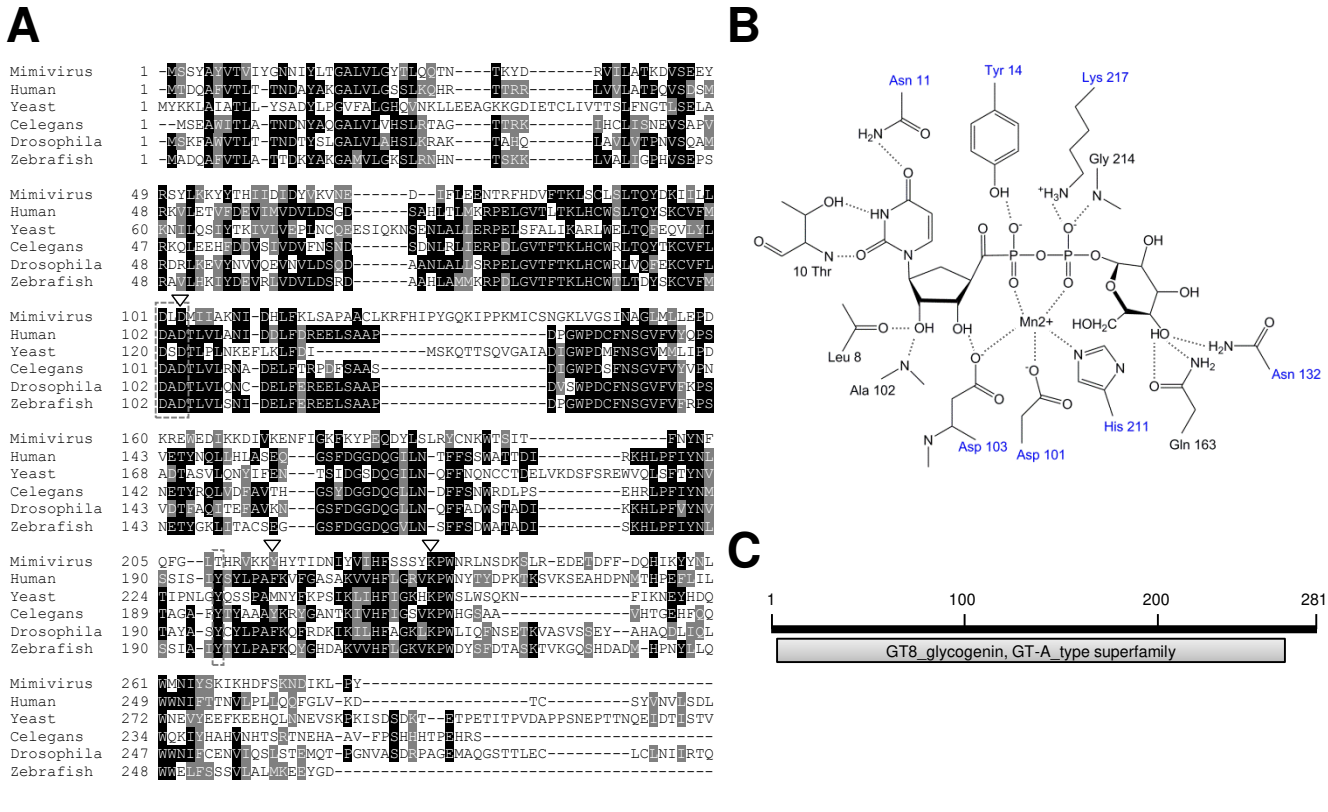

A
Mimivirus
Human
Yeast
Celegans
Drosophila
Zebrafish
B
Asn 11
Tyr 14
Lys 217
Gly 214
10 Thr
Leu 8
Ala 102
Mn2+
Asp 103
Asp 101
His 211
Gln 163
Asn 132
C
1
100
200
281
GT8_glycogenin, GT-A_type superfamily

Figure 2

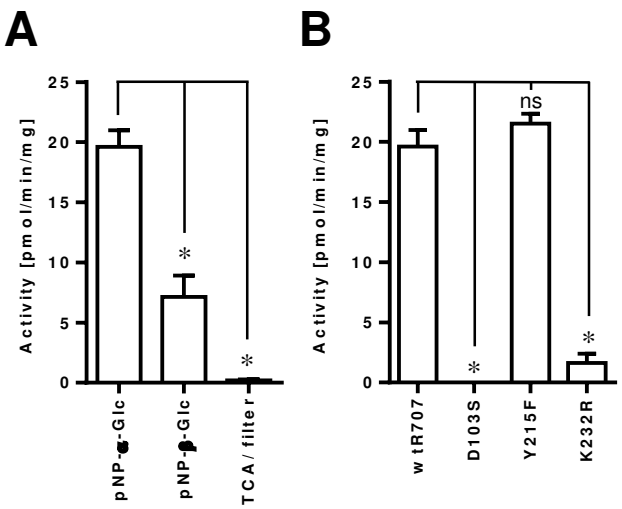

A
B
Activity [pmol/min/mg]
25
20
15
10
5
0
*
ns
pNP-α-Glc
pNP-β-Glc
TCA/filter
wtR707
D103S
Y215F
K232R

Figure 3

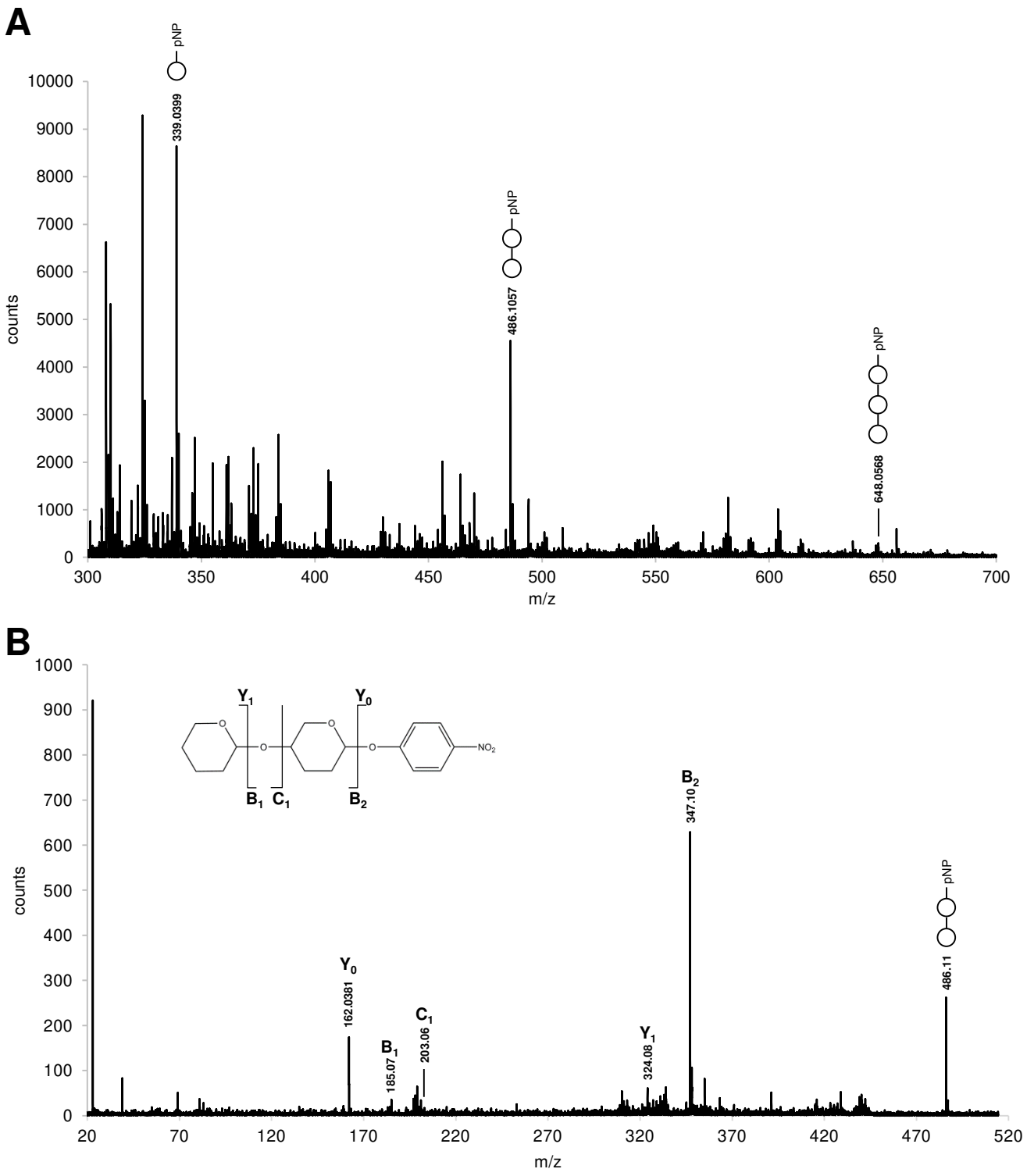

A
pNP
339.0399
pNP
486.1057
pNP
648.0568
counts
m/z
B
Y1
Y0
B1
C1
B2
NO2
B2
347.10
pNP
486.11
Y0
162.0381
C1
203.06
B1
185.07
Y1
324.08
counts
m/z

Figure 4

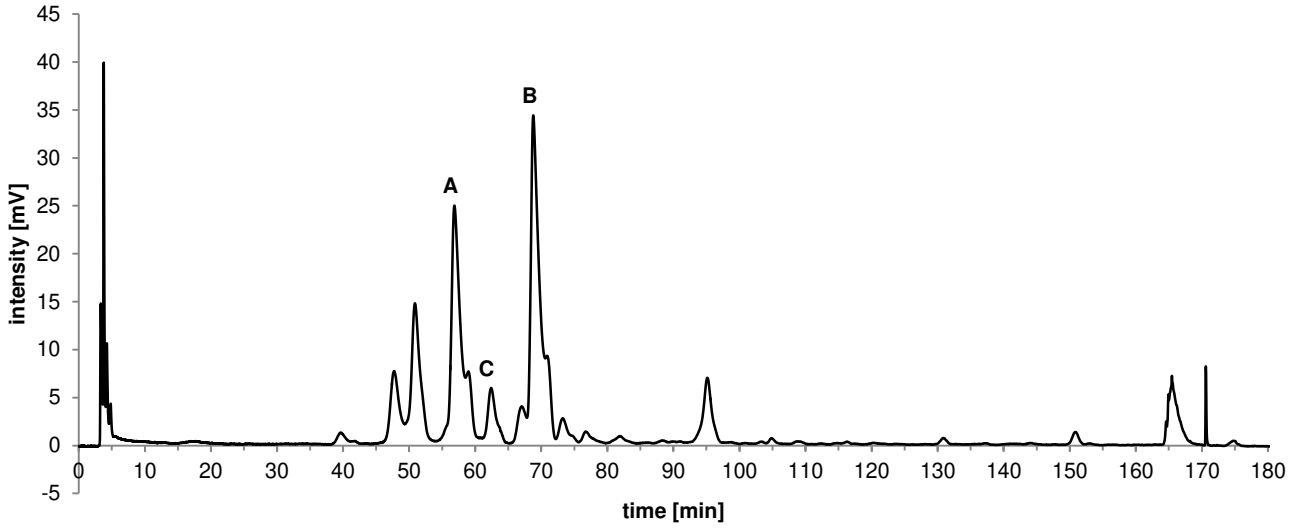

intensity [mV]
time [min]
A
B
C
45
40
35
30
25
20
15
10
5
0
-5
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180

Figure 5

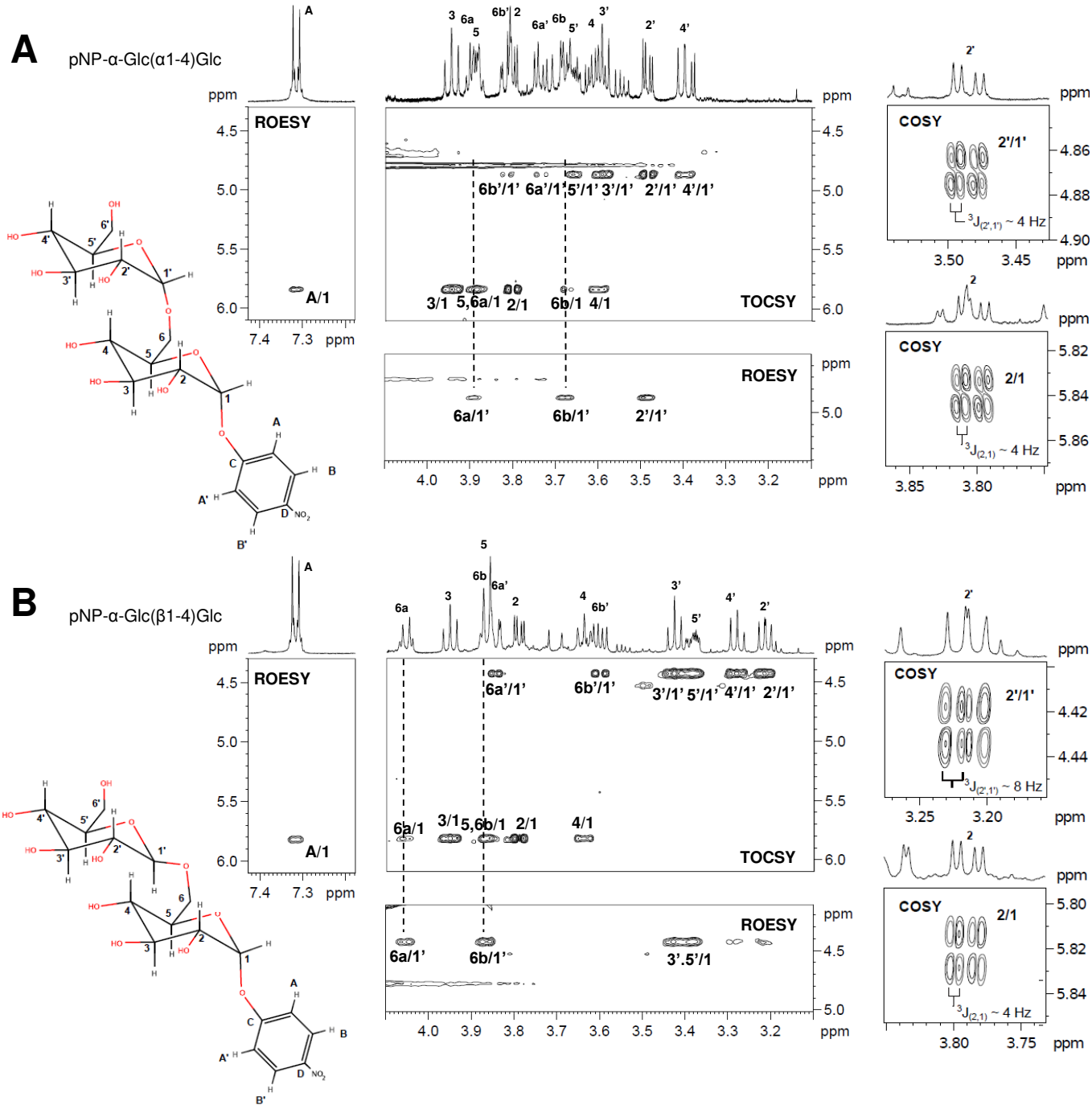
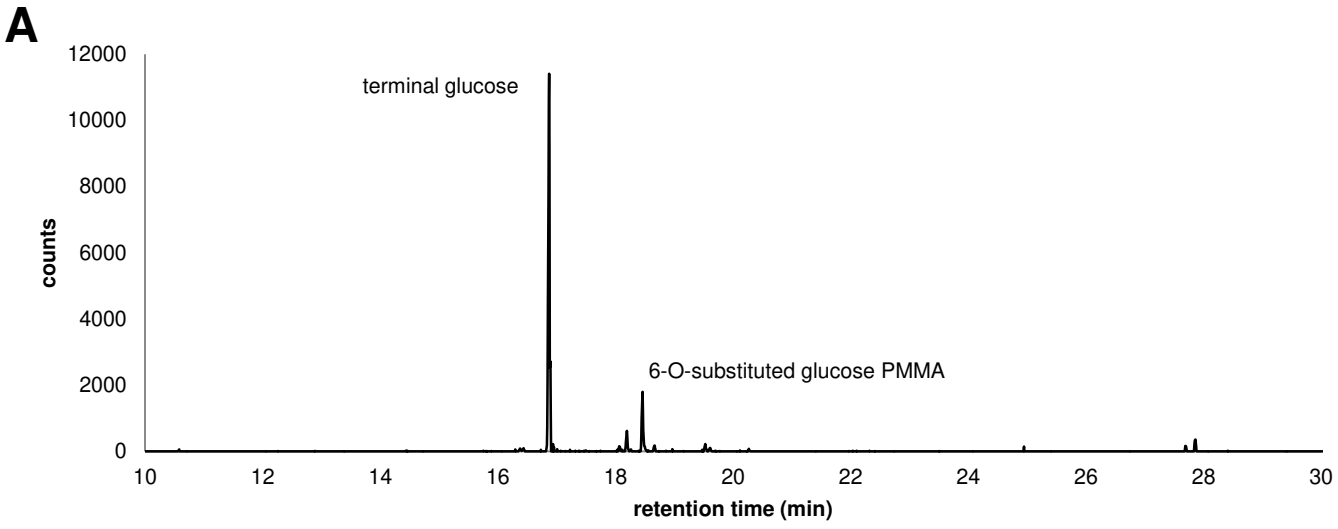
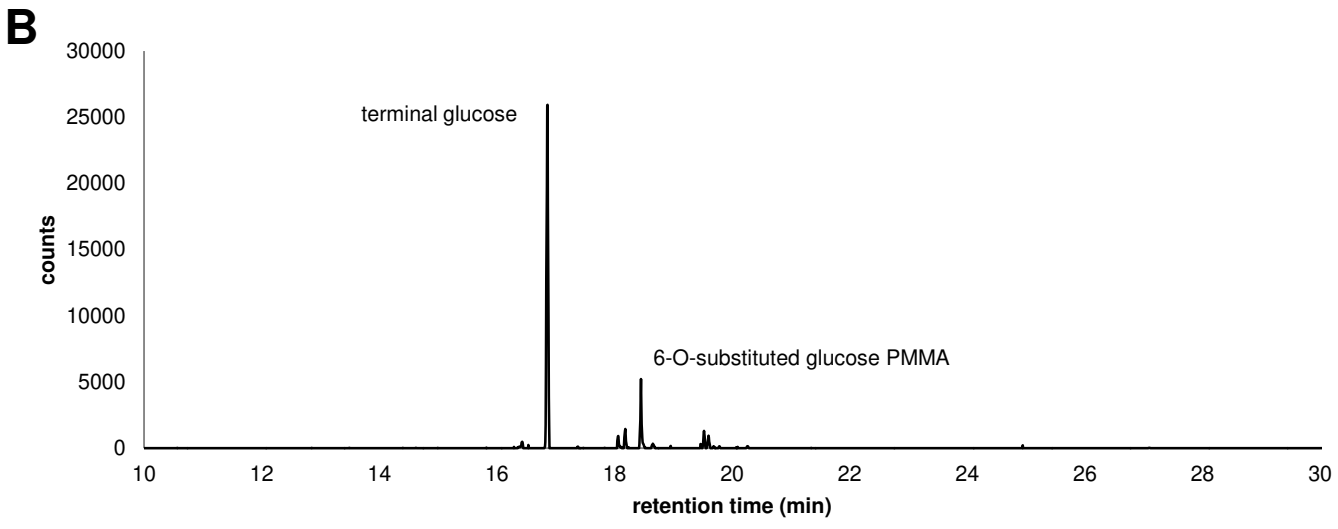

A
pNP-α-Glc(α1-4)Glc
ROESY
A/1
TOCSY
3/1 5,6a/1 2/1 6b/1 4/1
6b'/1' 6a'/1' 5'/1' 3'/1' 2'/1' 4'/1'
ROESY
6a/1' 6b/1' 2'/1'
COSY
2'/1'
$^3J_{(2',1')}$ ~ 4 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz
B
pNP-α-Glc(β1-4)Glc
ROESY
A/1
TOCSY
6a/1 3/1 5,6b/1 2/1 4/1
6a'/1' 6b'/1' 3'/1' 5'/1' 4'/1' 2'/1'
ROESY
6a/1' 6b/1' 3'.5'/1
COSY
2'/1'
$^3J_{(2',1')}$ ~ 8 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz

Figure 6


A
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)


B
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)

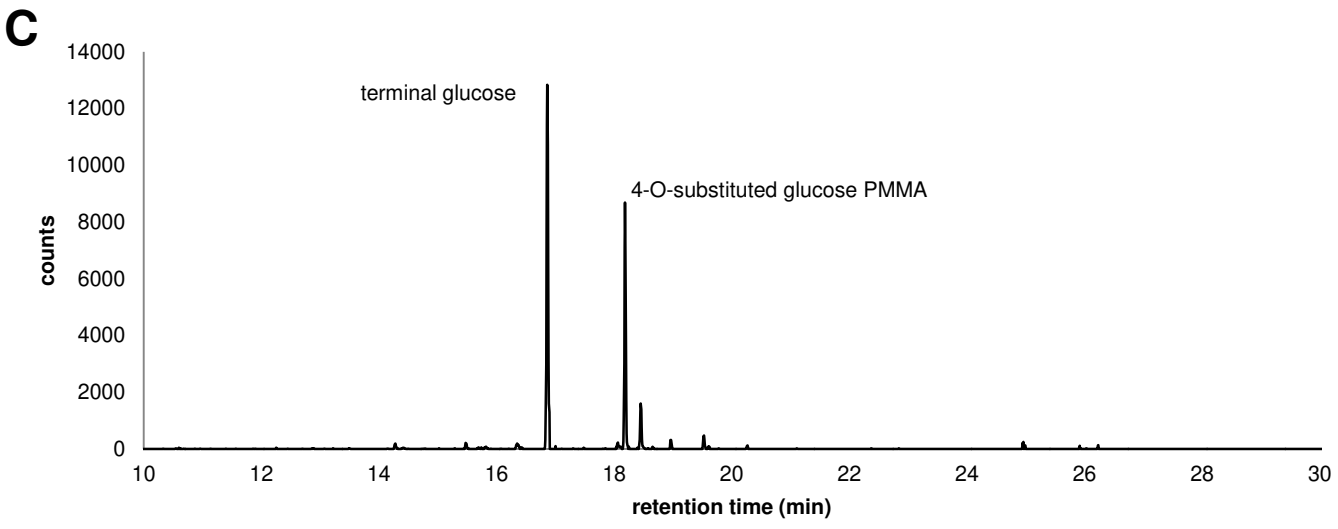
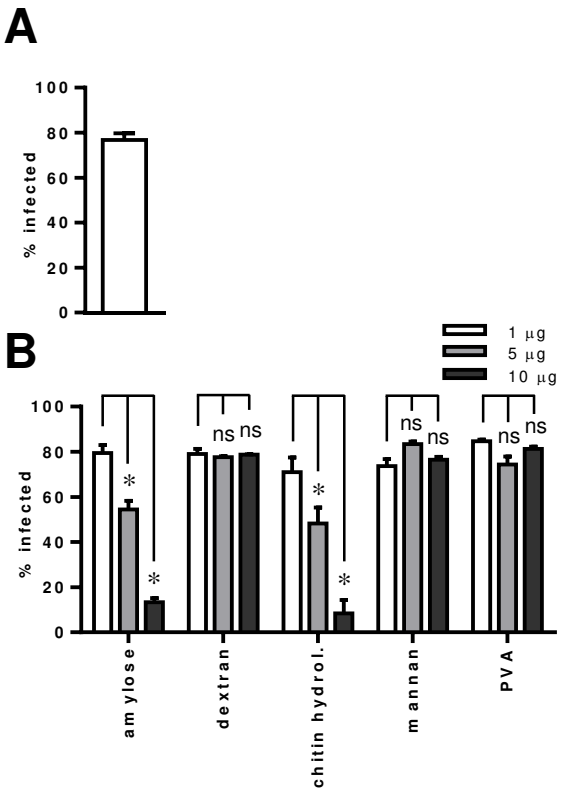
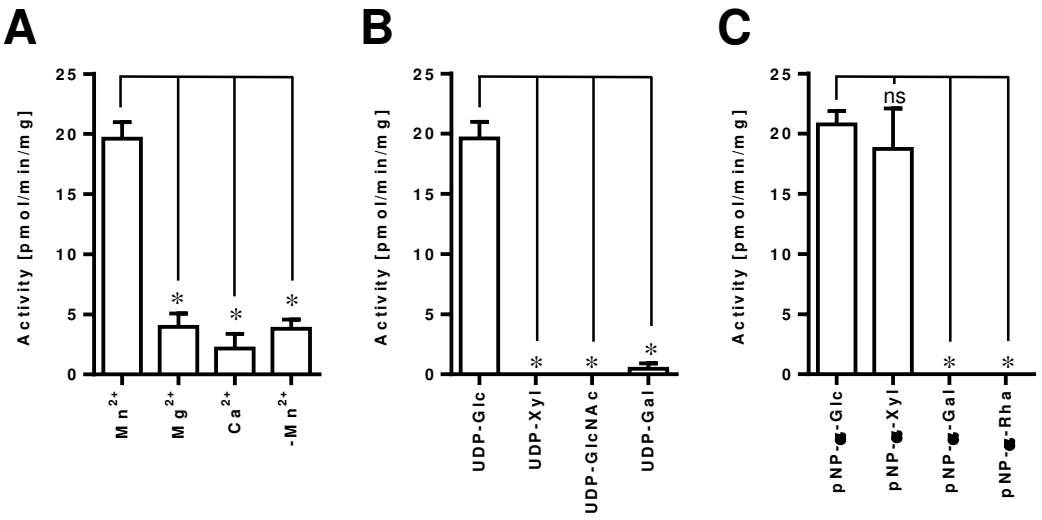

C
terminal glucose
4-O-substituted glucose PMMA
counts
retention time (min)

Figure 7

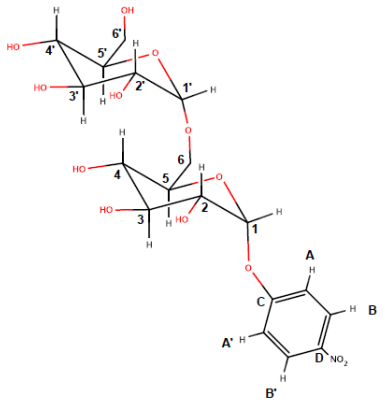

A
100
80
60
40
20
0
% infected
B
1 µg
5 µg
10 µg
100
80
60
40
20
0
% infected
*
*
*
ns
ns
*
ns
ns
ns
ns
amylose
dextran
chitin hydrol.
mannan
PVA

Supplementary Figure 1



Supplementary Figure 2

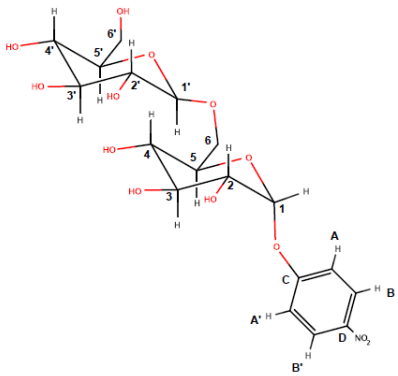
A



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.84	3.80	3.95	3.60	3.88	3.90	3.68	4.87	3.48
δ(^{13}C) / ppm	99.1	73.4	75.7	71.9	74.1	68.0	68.0	100.2	73.9

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.59	3.40	3.66	3.74	3.82	7.31	8.28		
δ(^{13}C) / ppm	75.6	72.0	74.3	63.0	63.0	119.4	128.6	164.0	145.1

B



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.82	3.79	3.95	3.64	3.86	4.06	3.87	4.42	3.21
δ(^{13}C) / ppm	99.2	73.4	75.4	71.5	74.6	70.7	70.7	105.2	75.7

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.43	3.28	3.39	3.84	3.60	7.32	8.28		
δ(^{13}C) / ppm	78.3	72.2	78.5	63.4	63.4	119.5	128.6	164.0	145.1

Supplementary Figure 3

HSQC signals Peak C

δ(^{1}H) / ppm	3.75	3.74	3.82	3.4	3.8	3.78	3.58	3.68	3.66
δ(^{13}C) / ppm	62.9	63.1	63.1	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.21	3.76	5.8	5.42	7.28	8.23
δ(^{13}C) / ppm	76	79.1	99.1	102.4	119.4	128.7

HSQC signals pNP-maltose

δ(^{1}H) / ppm	3.76	3.74	3.83	3.41	3.81	3.8	3.57	3.68	3.67
δ(^{13}C) / ppm	62.9	63.2	63.2	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.22	3.77	5.82	5.43	7.31	8.27
δ(^{13}C) / ppm	76.1	79	99.1	102.4	119.5	128.7

Supplementary Figure 4

Figure 1

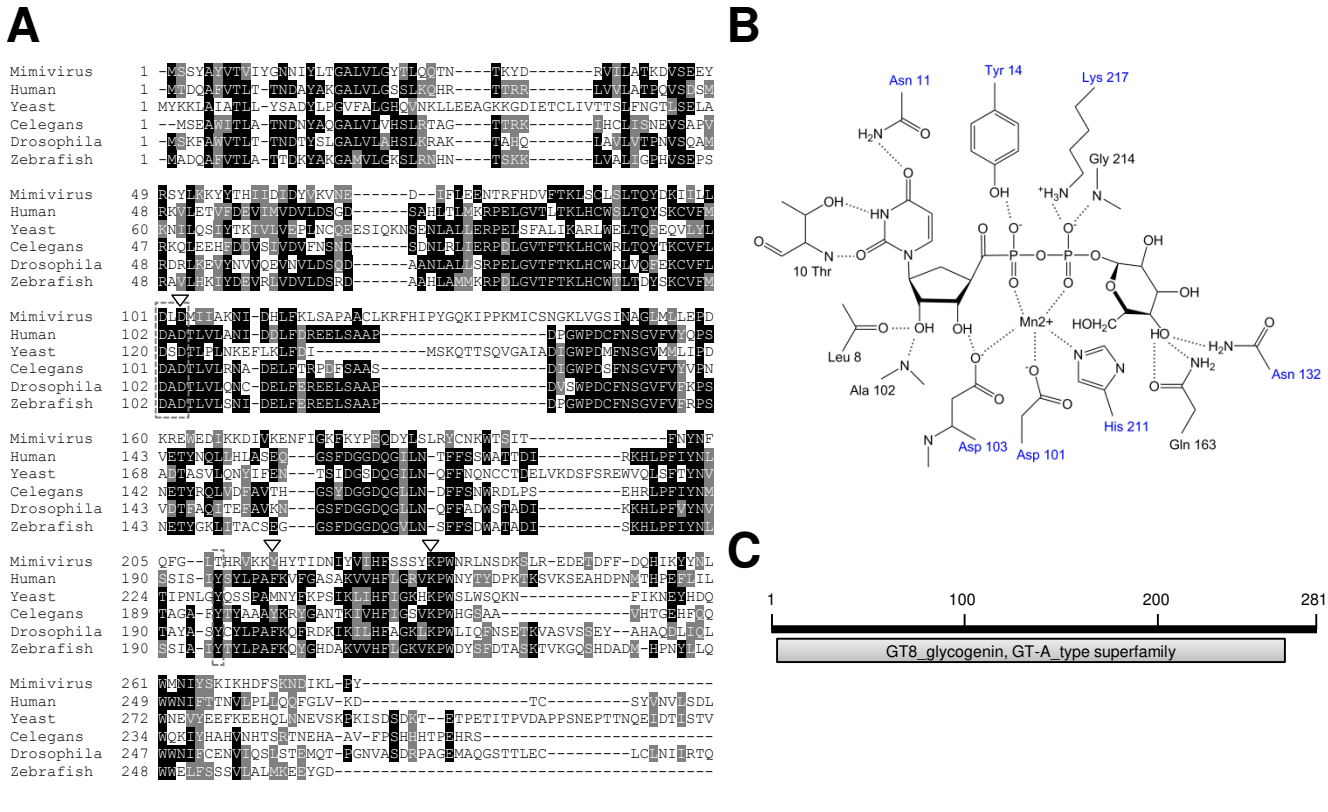

A
Mimivirus
Human
Yeast
Celegans
Drosophila
Zebrafish
B
Asn 11
Tyr 14
Lys 217
Gly 214
10 Thr
Leu 8
Ala 102
Mn2+
Asp 103
Asp 101
His 211
Gln 163
Asn 132
C
1
100
200
281
GT8_glycogenin, GT-A_type superfamily

Figure 2

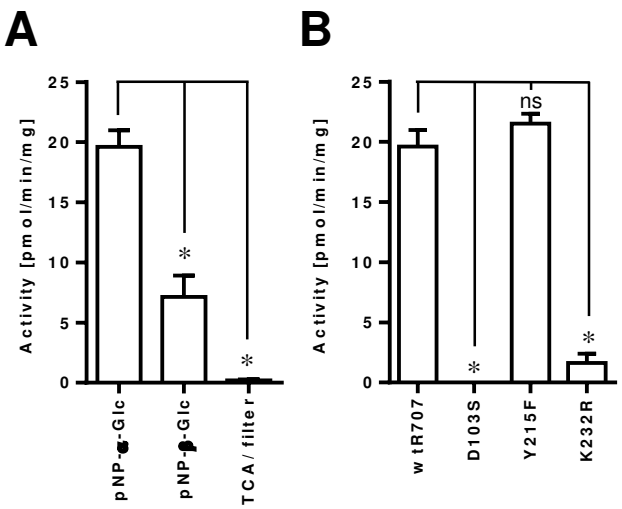

A
B
Activity [pmol/min/mg]
25
20
15
10
5
0
*
ns
pNP-α-Glc
pNP-β-Glc
TCA/filter
wtR707
D103S
Y215F
K232R

Figure 3

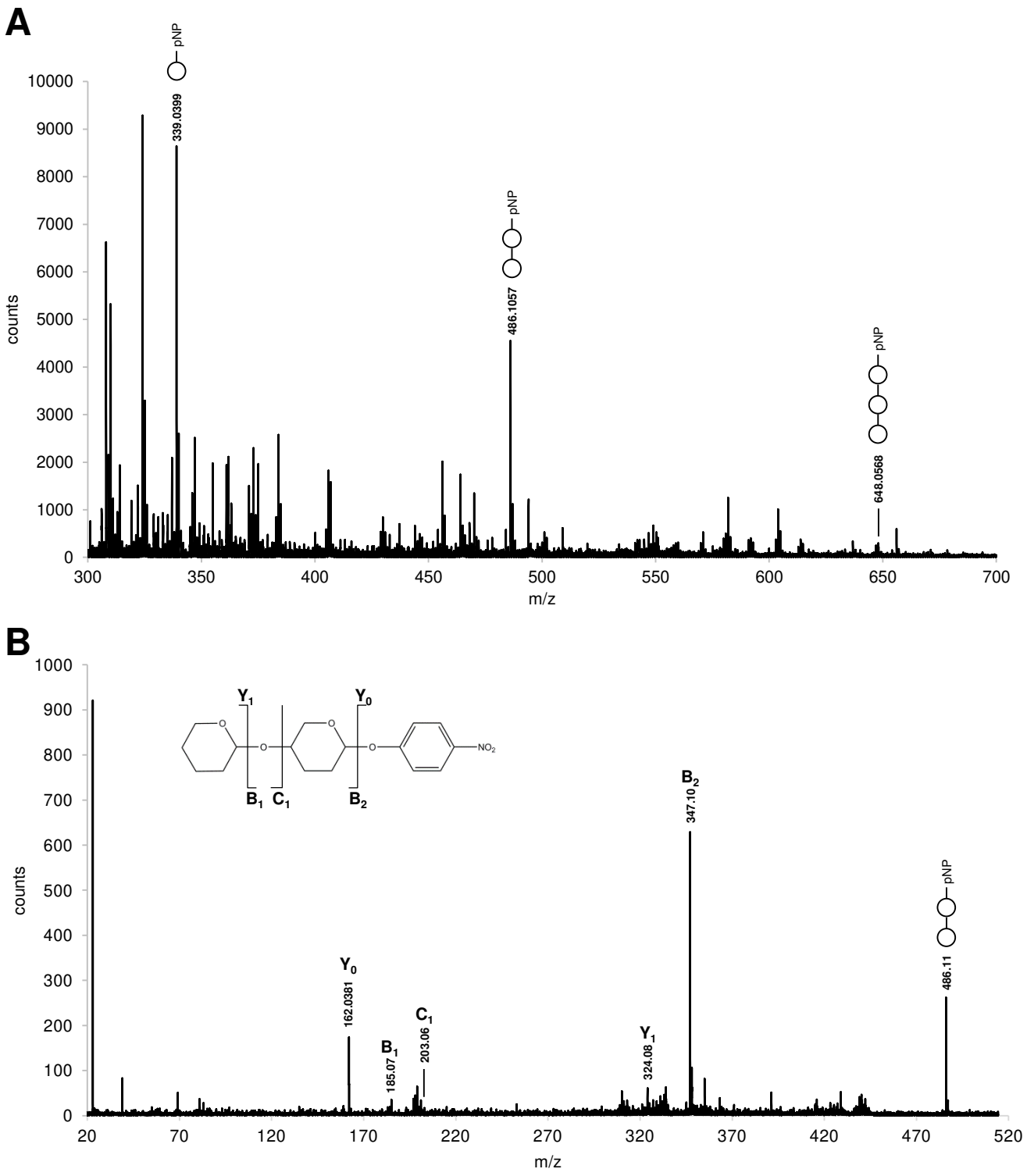

A
pNP
339.0399
pNP
486.1057
pNP
648.0568
counts
m/z
B
Y1
Y0
B1
C1
B2
NO2
B2
347.10
pNP
486.11
Y0
162.0381
C1
203.06
B1
185.07
Y1
324.08
counts
m/z

Figure 4

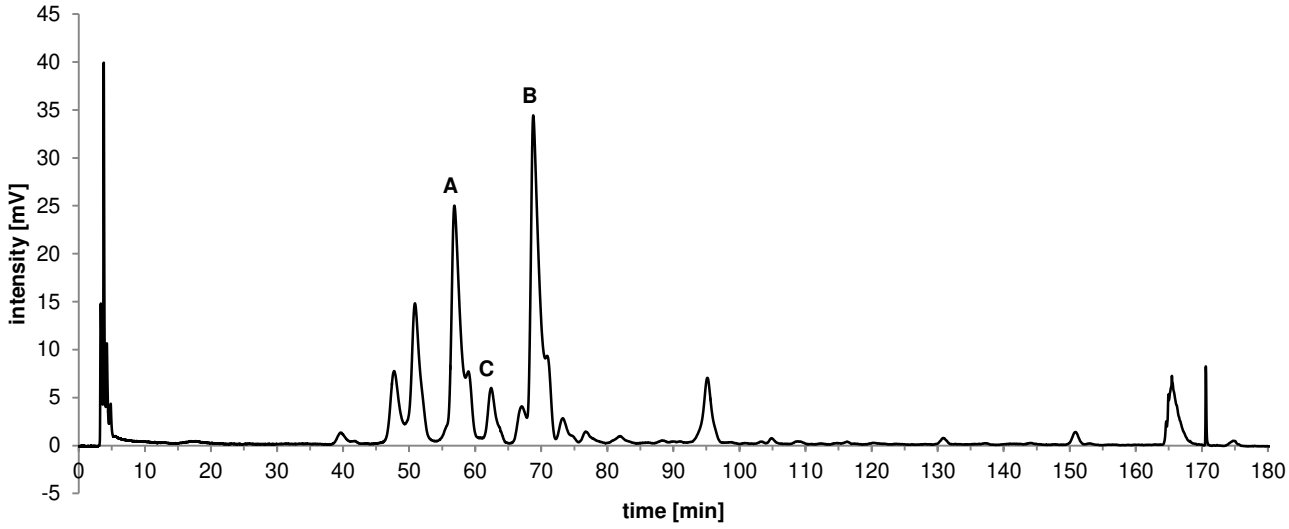

intensity [mV]
time [min]
A
B
C
45
40
35
30
25
20
15
10
5
0
-5
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180

Figure 5

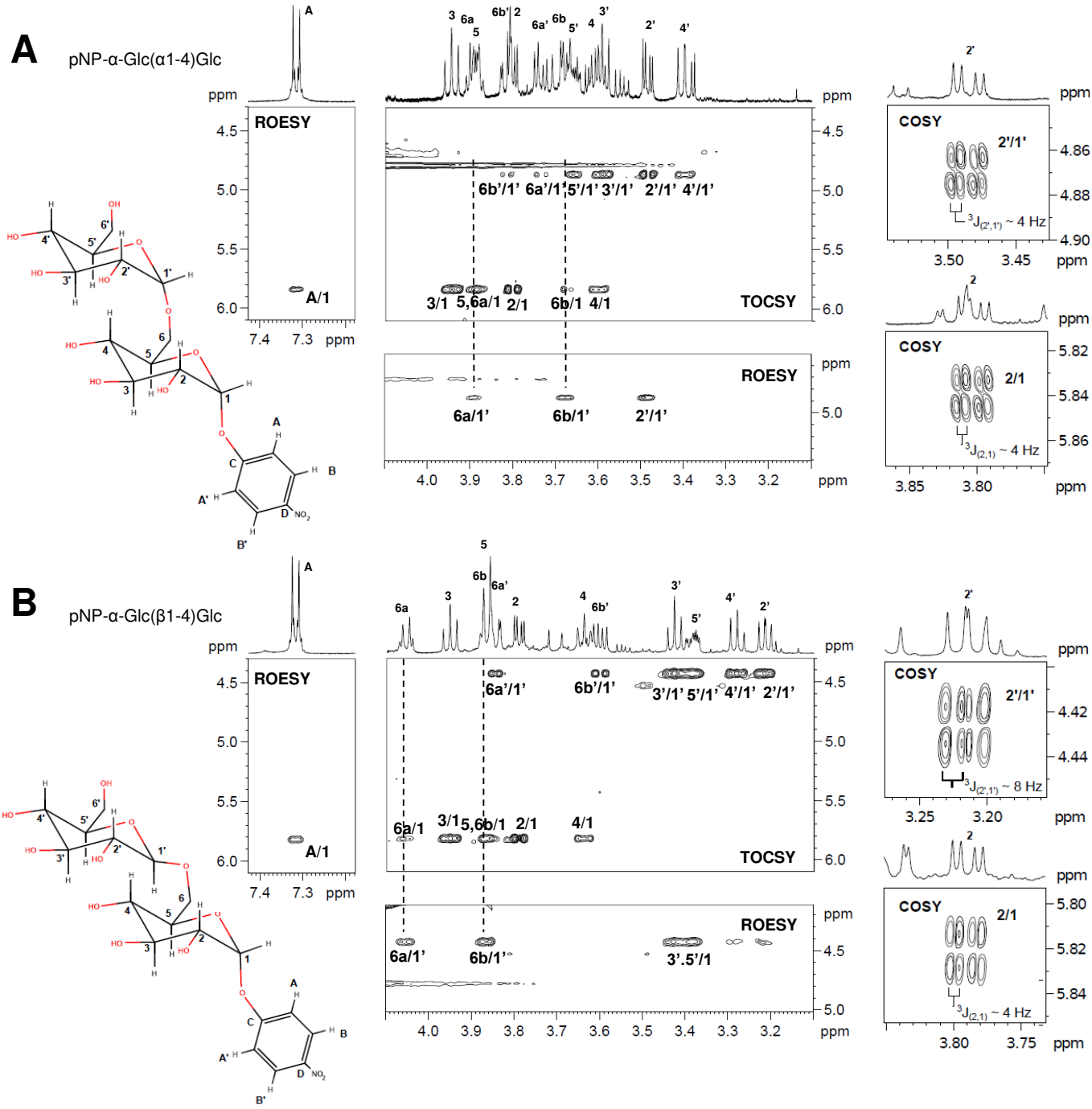
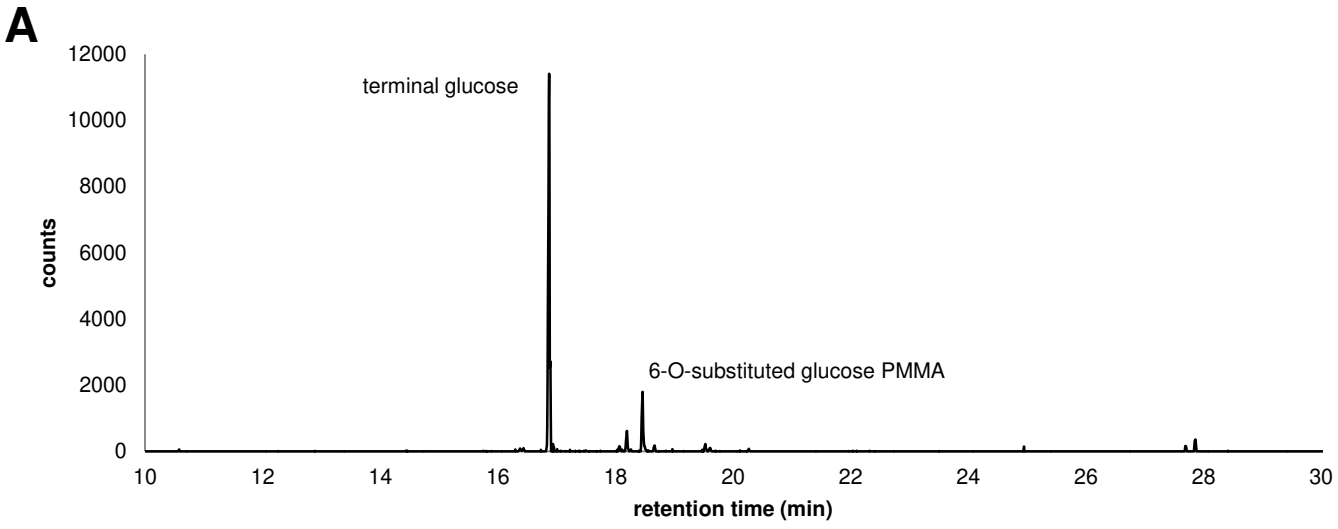
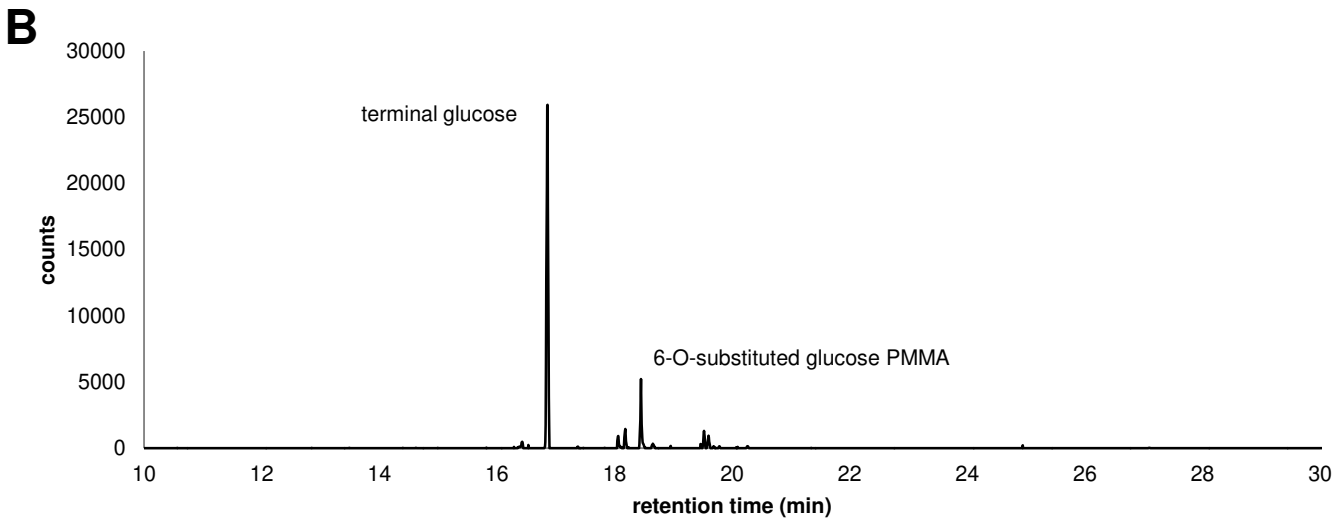

A
pNP-α-Glc(α1-4)Glc
ROESY
A/1
TOCSY
3/1 5,6a/1 2/1 6b/1 4/1
6b'/1' 6a'/1' 5'/1' 3'/1' 2'/1' 4'/1'
ROESY
6a/1' 6b/1' 2'/1'
COSY
2'/1'
$^3J_{(2',1')}$ ~ 4 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz
B
pNP-α-Glc(β1-4)Glc
ROESY
A/1
TOCSY
6a/1 3/1 5,6b/1 2/1 4/1
6a'/1' 6b'/1' 3'/1' 5'/1' 4'/1' 2'/1'
ROESY
6a/1' 6b/1' 3'.5'/1
COSY
2'/1'
$^3J_{(2',1')}$ ~ 8 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz

Figure 6


A
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)


B
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)

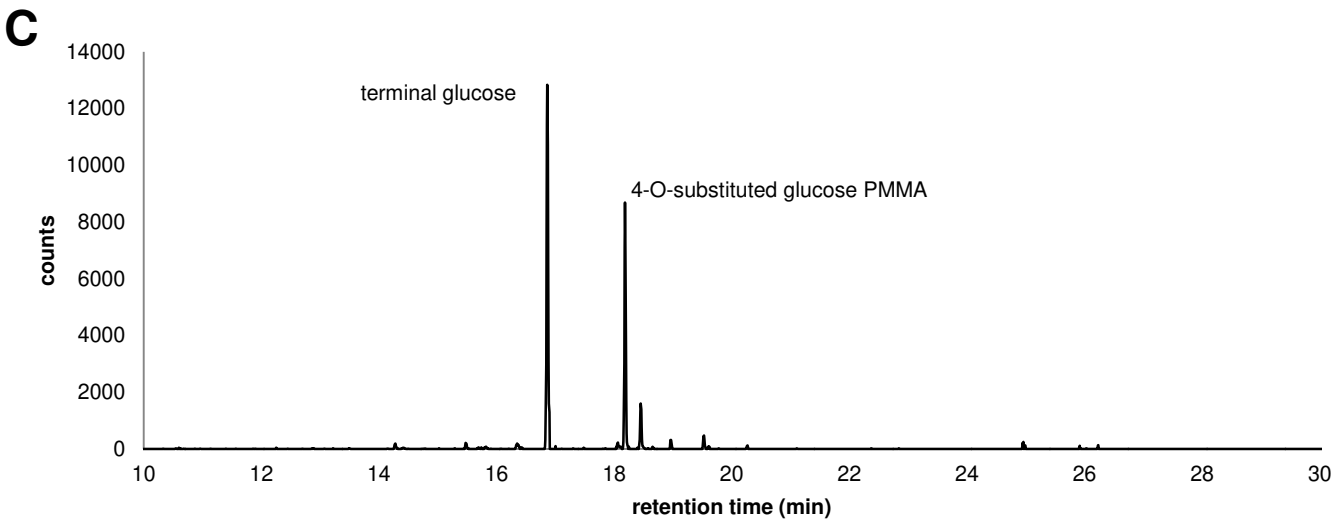
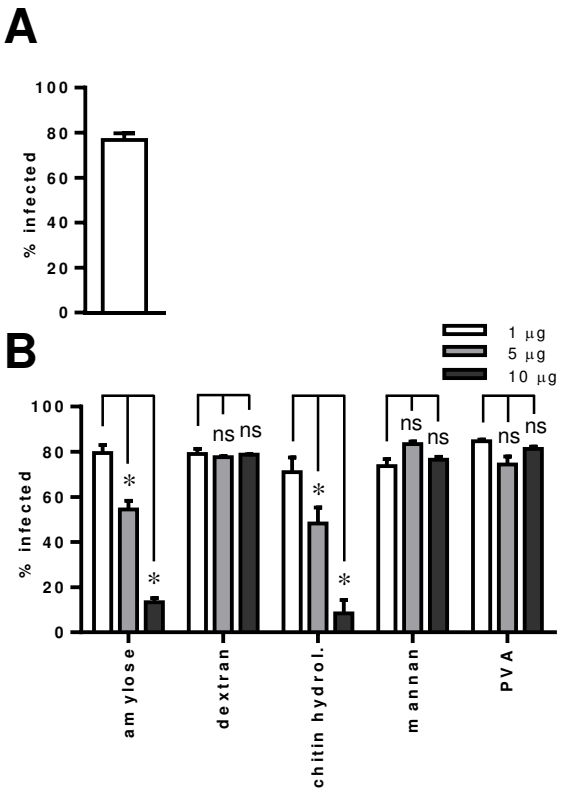
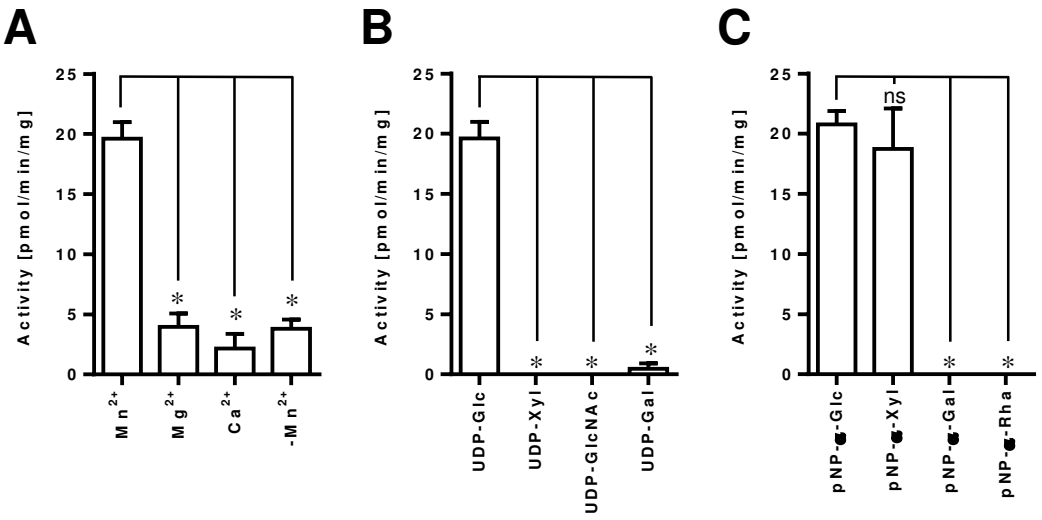

C
terminal glucose
4-O-substituted glucose PMMA
counts
retention time (min)

Figure 7

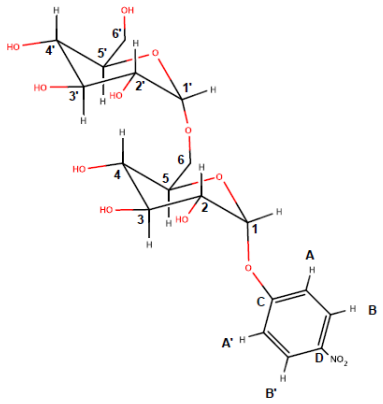

A
100
80
60
40
20
0
% infected
B
1 µg
5 µg
10 µg
100
80
60
40
20
0
% infected
*
*
*
ns
ns
*
ns
ns
ns
ns
amylose
dextran
chitin hydrol.
mannan
PVA

Supplementary Figure 1



Supplementary Figure 2

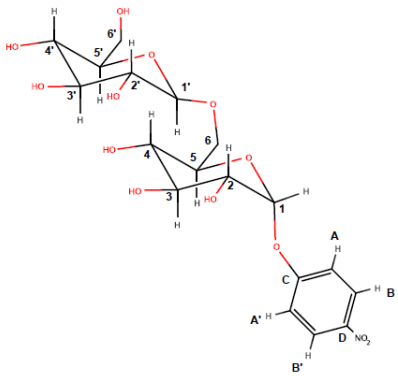
A



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.84	3.80	3.95	3.60	3.88	3.90	3.68	4.87	3.48
δ(^{13}C) / ppm	99.1	73.4	75.7	71.9	74.1	68.0	68.0	100.2	73.9

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.59	3.40	3.66	3.74	3.82	7.31	8.28		
δ(^{13}C) / ppm	75.6	72.0	74.3	63.0	63.0	119.4	128.6	164.0	145.1

B



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.82	3.79	3.95	3.64	3.86	4.06	3.87	4.42	3.21
δ(^{13}C) / ppm	99.2	73.4	75.4	71.5	74.6	70.7	70.7	105.2	75.7

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.43	3.28	3.39	3.84	3.60	7.32	8.28		
δ(^{13}C) / ppm	78.3	72.2	78.5	63.4	63.4	119.5	128.6	164.0	145.1

Supplementary Figure 3

HSQC signals Peak C

δ(^{1}H) / ppm	3.75	3.74	3.82	3.4	3.8	3.78	3.58	3.68	3.66
δ(^{13}C) / ppm	62.9	63.1	63.1	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.21	3.76	5.8	5.42	7.28	8.23
δ(^{13}C) / ppm	76	79.1	99.1	102.4	119.4	128.7

HSQC signals pNP-maltose

δ(^{1}H) / ppm	3.76	3.74	3.83	3.41	3.81	3.8	3.57	3.68	3.67
δ(^{13}C) / ppm	62.9	63.2	63.2	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.22	3.77	5.82	5.43	7.31	8.27
δ(^{13}C) / ppm	76.1	79	99.1	102.4	119.5	128.7

Supplementary Figure 4

Figure 1

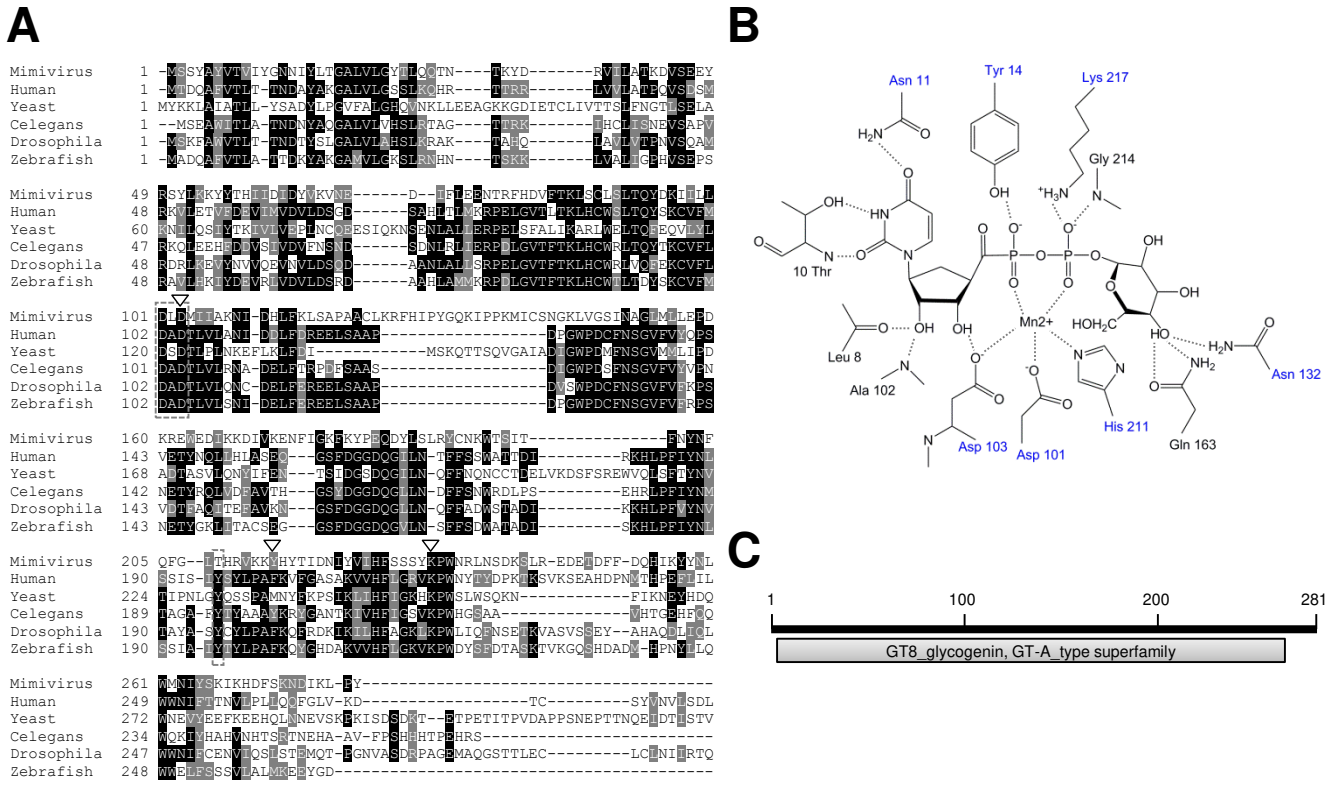

A
Mimivirus
Human
Yeast
Celegans
Drosophila
Zebrafish
B
Asn 11
Tyr 14
Lys 217
Gly 214
10 Thr
Leu 8
Ala 102
Mn2+
Asp 103
Asp 101
His 211
Gln 163
Asn 132
C
1
100
200
281
GT8_glycogenin, GT-A_type superfamily

Figure 2

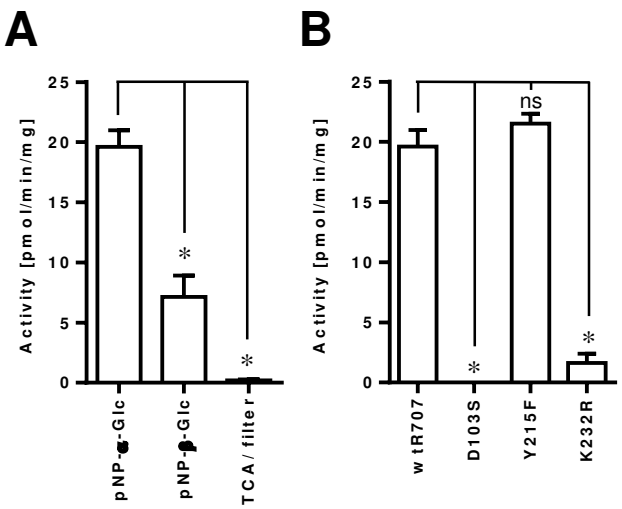

A
B
Activity [pmol/min/mg]
25
20
15
10
5
0
*
ns
pNP-α-Glc
pNP-β-Glc
TCA/filter
wtR707
D103S
Y215F
K232R

Figure 3

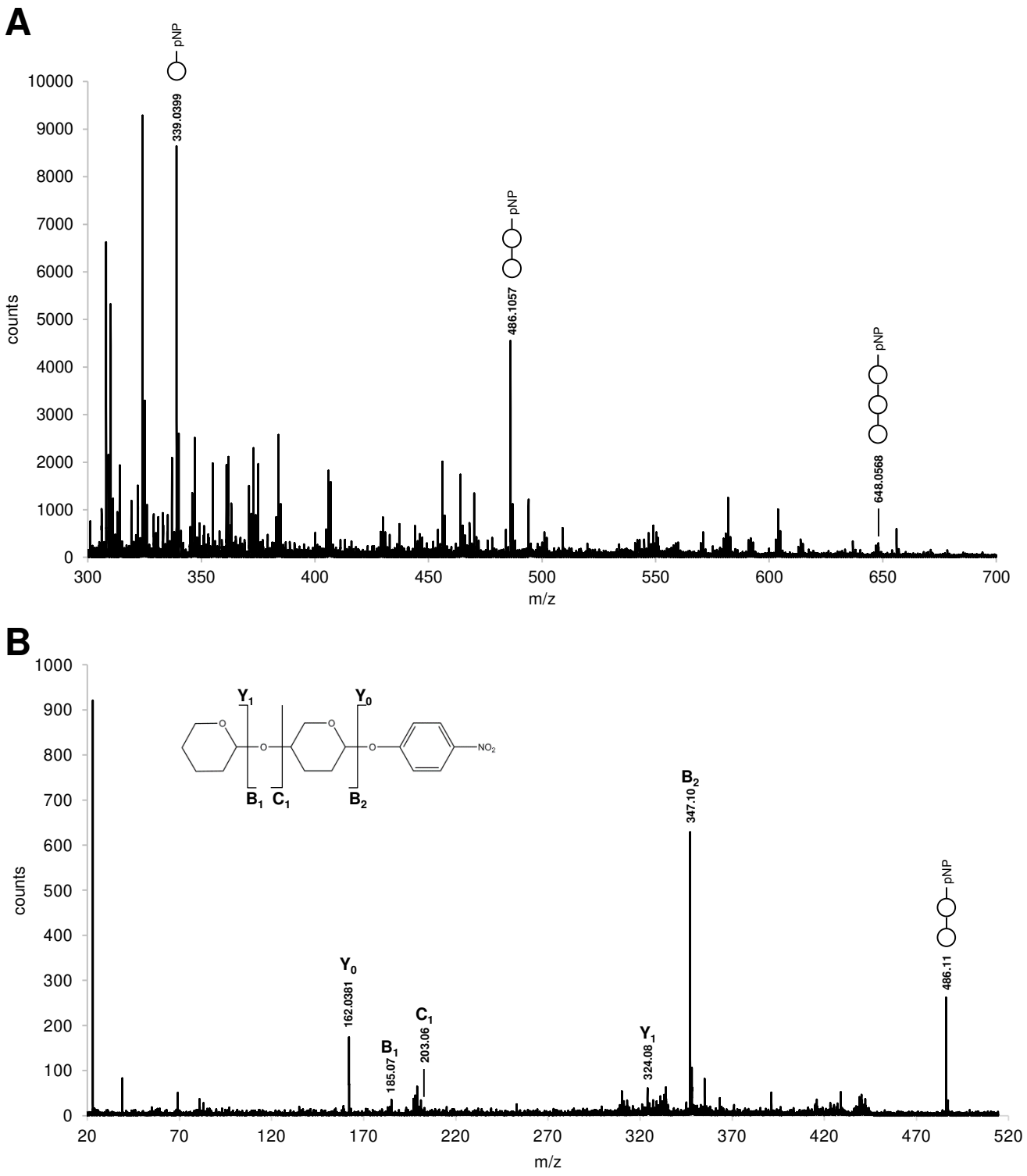

A
pNP
339.0399
pNP
486.1057
pNP
648.0568
counts
m/z
B
Y1
Y0
B1
C1
B2
NO2
B2
347.10
pNP
486.11
Y0
162.0381
C1
203.06
B1
185.07
Y1
324.08
counts
m/z

Figure 4

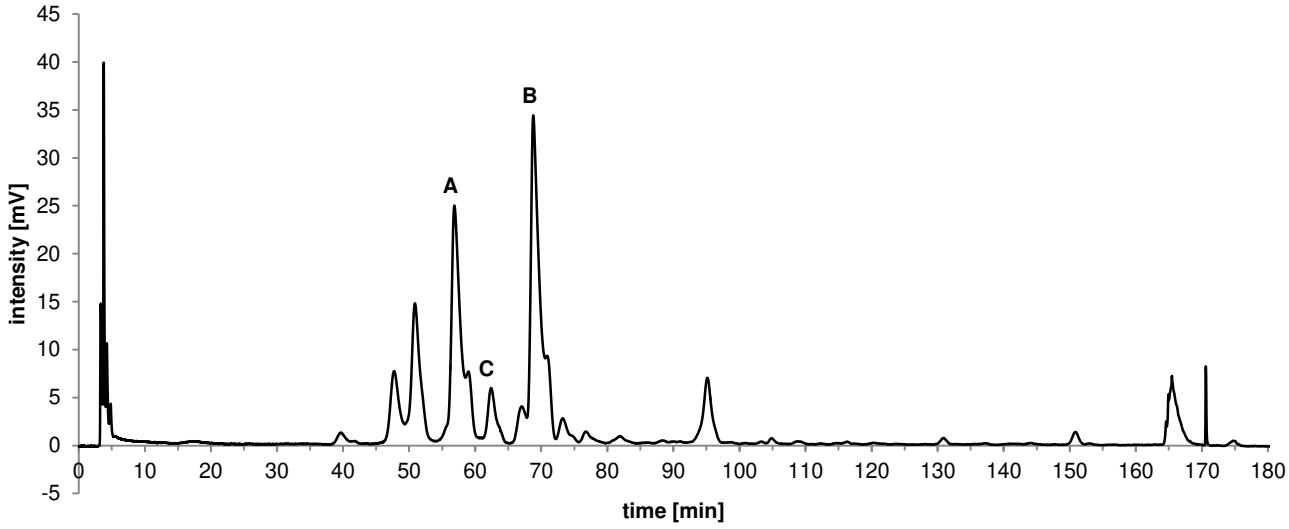

intensity [mV]
time [min]
A
B
C
45
40
35
30
25
20
15
10
5
0
-5
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180

Figure 5

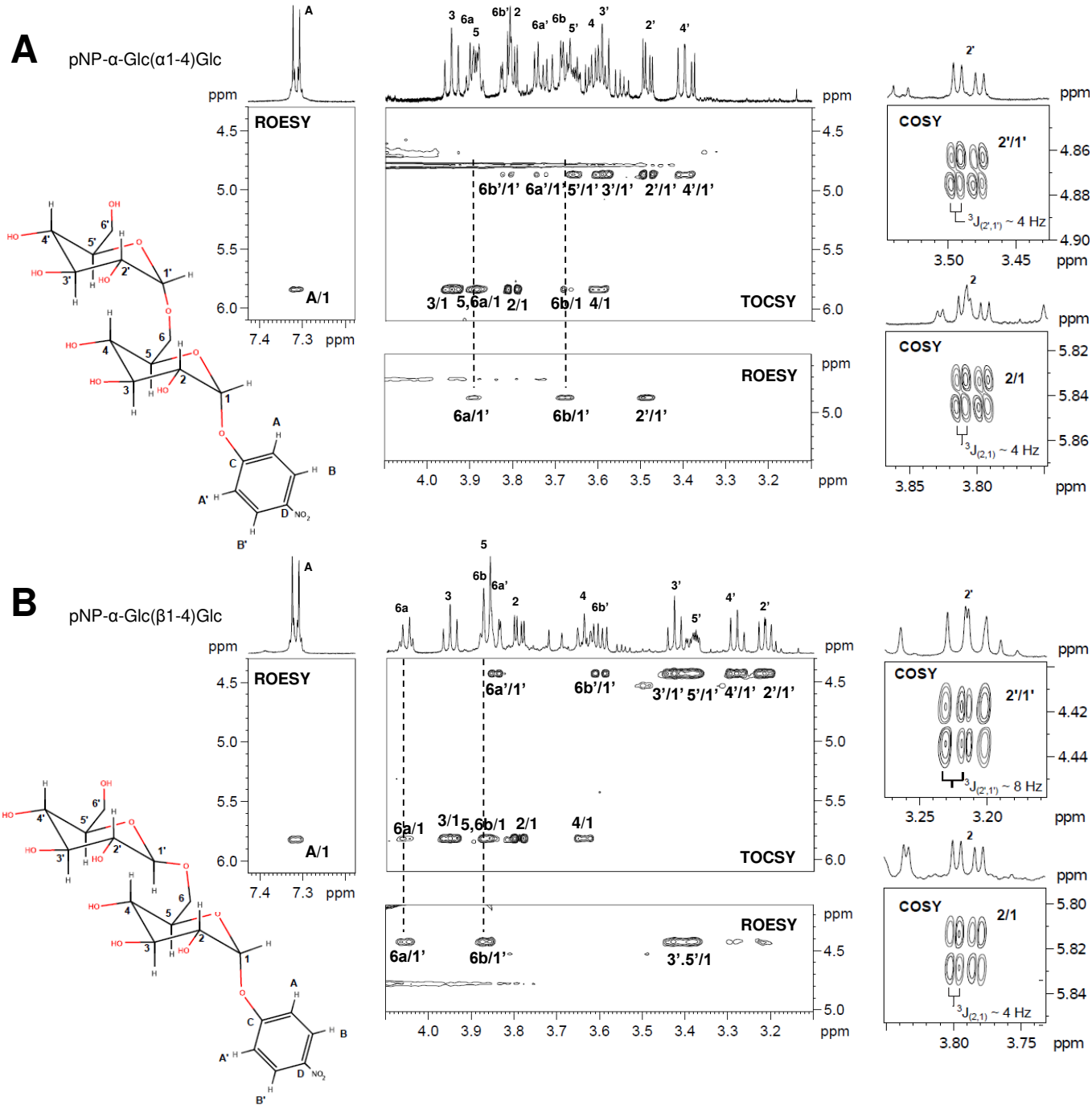
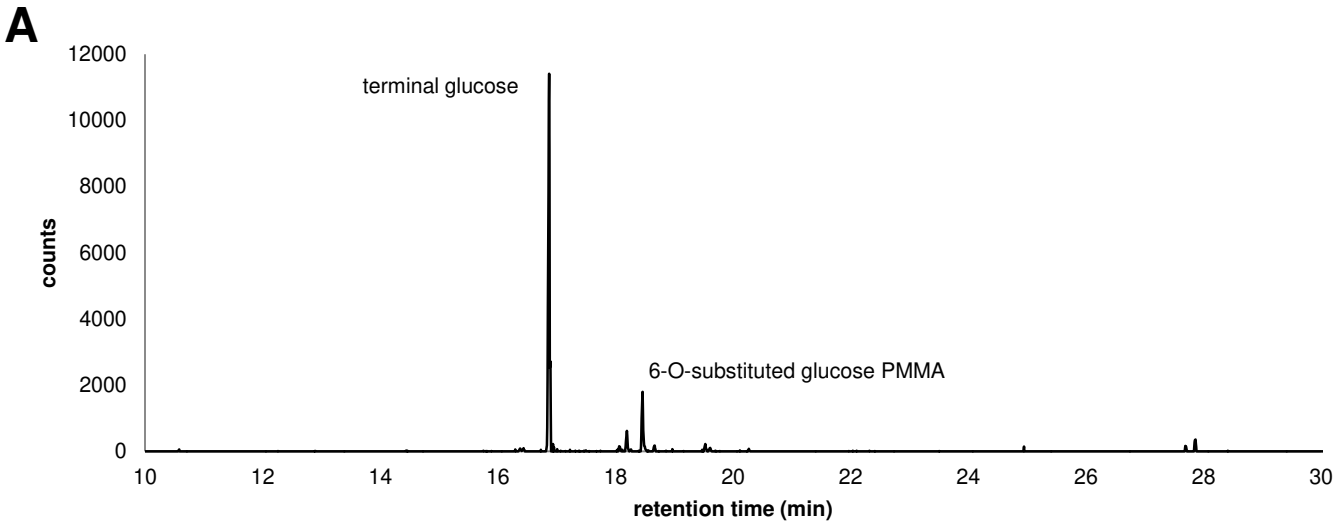
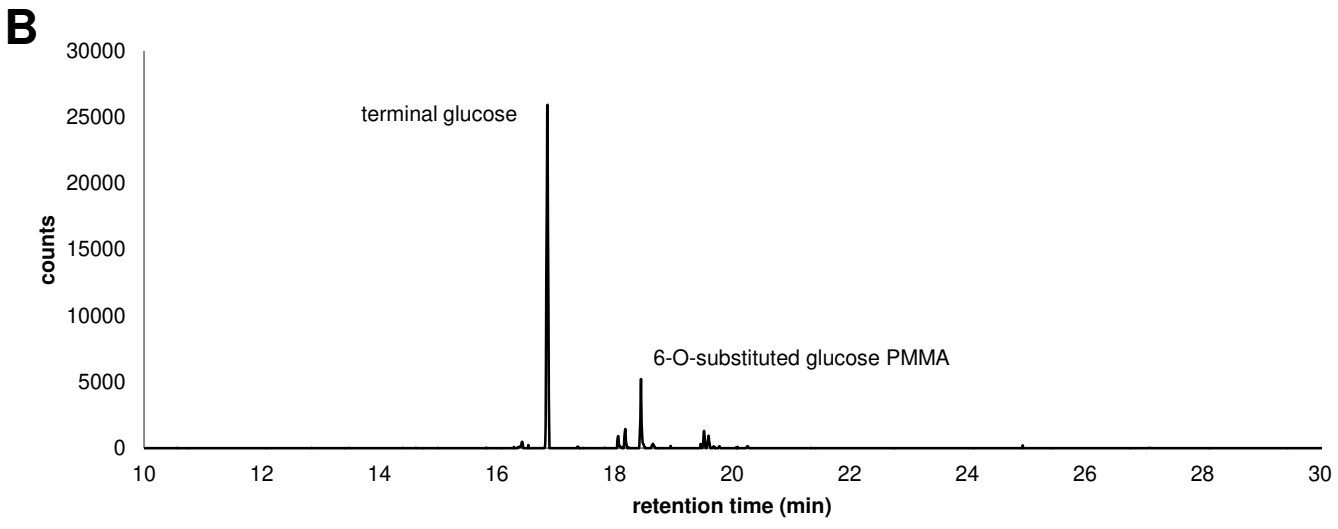

A
pNP-α-Glc(α1-4)Glc
ROESY
A/1
TOCSY
3/1 5,6a/1 2/1 6b/1 4/1
6b'/1' 6a'/1' 5'/1' 3'/1' 2'/1' 4'/1'
ROESY
6a/1' 6b/1' 2'/1'
COSY
2'/1'
$^3J_{(2',1')}$ ~ 4 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz
B
pNP-α-Glc(β1-4)Glc
ROESY
A/1
TOCSY
6a/1 3/1 5,6b/1 2/1 4/1
6a'/1' 6b'/1' 3'/1' 5'/1' 4'/1' 2'/1'
ROESY
6a/1' 6b/1' 3'.5'/1
COSY
2'/1'
$^3J_{(2',1')}$ ~ 8 Hz
COSY
2/1
$^3J_{(2,1)}$ ~ 4 Hz

Figure 6


A
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)


B
terminal glucose
6-O-substituted glucose PMMA
counts
retention time (min)

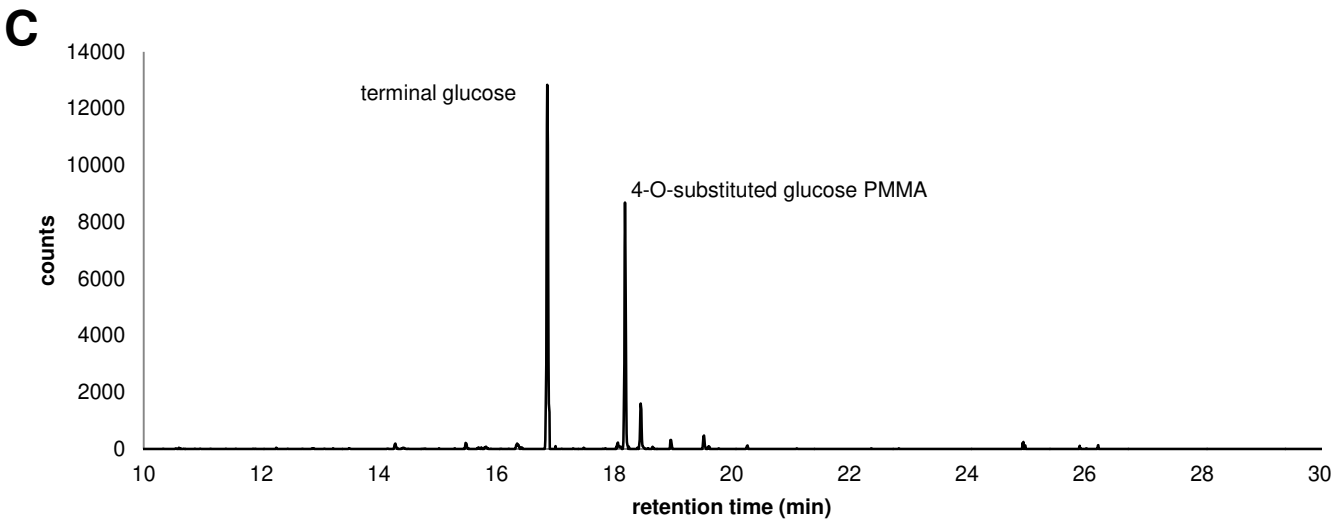
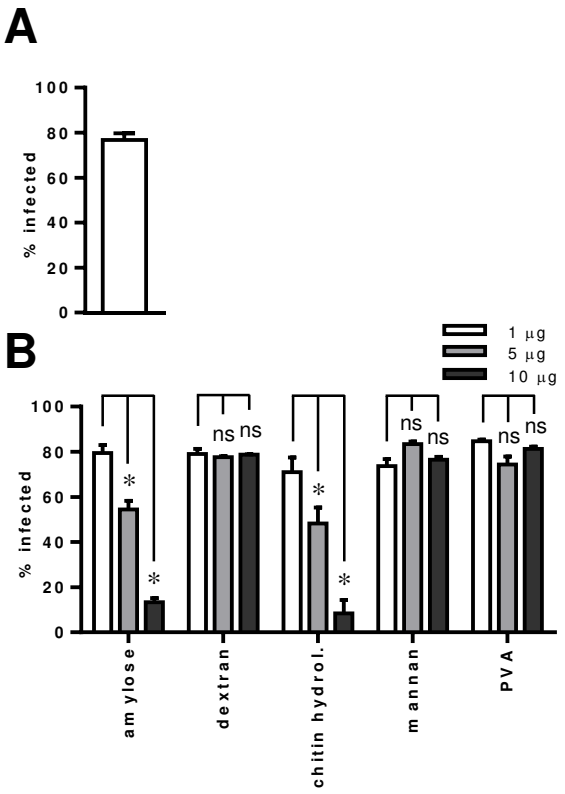
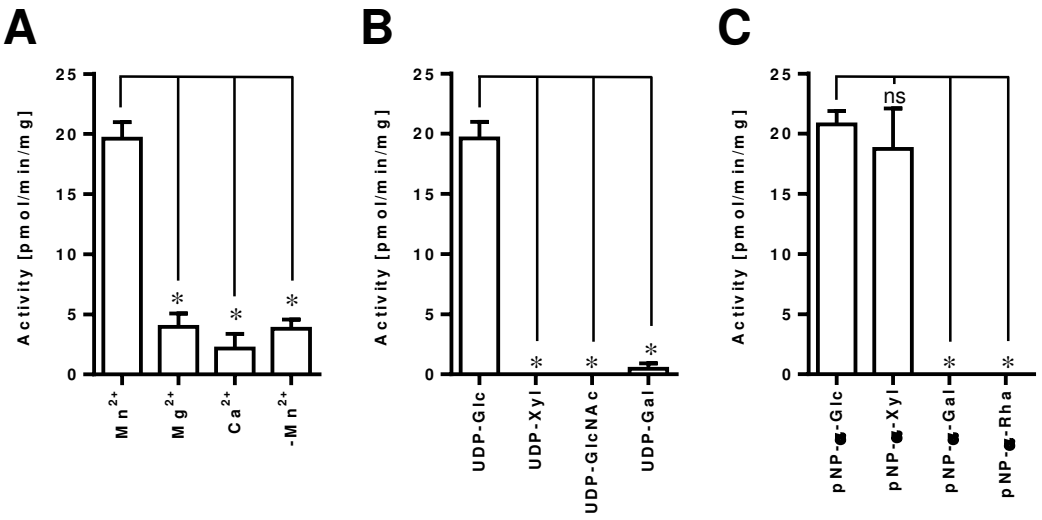

C
terminal glucose
4-O-substituted glucose PMMA
counts
retention time (min)

Figure 7

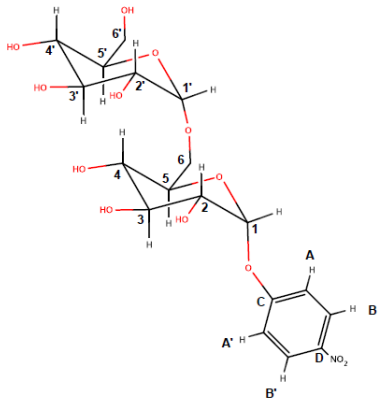

A
100
80
60
40
20
0
% infected
B
1 µg
5 µg
10 µg
100
80
60
40
20
0
% infected
*
*
*
ns
ns
*
ns
ns
ns
ns
amylose
dextran
chitin hydrol.
mannan
PVA

Supplementary Figure 1



Supplementary Figure 2

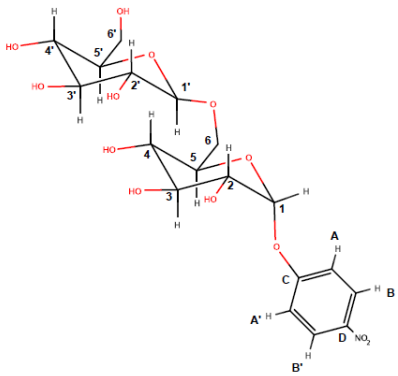
A



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.84	3.80	3.95	3.60	3.88	3.90	3.68	4.87	3.48
δ(^{13}C) / ppm	99.1	73.4	75.7	71.9	74.1	68.0	68.0	100.2	73.9

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.59	3.40	3.66	3.74	3.82	7.31	8.28		
δ(^{13}C) / ppm	75.6	72.0	74.3	63.0	63.0	119.4	128.6	164.0	145.1

B



Position	1	2	3	4	5	6a	6b	1’	2’
δ(^{1}H) / ppm	5.82	3.79	3.95	3.64	3.86	4.06	3.87	4.42	3.21
δ(^{13}C) / ppm	99.2	73.4	75.4	71.5	74.6	70.7	70.7	105.2	75.7

Position	3’	4’	5’	6a’	6b’	A,A’	B,B’	C	D
δ(^{1}H) / ppm	3.43	3.28	3.39	3.84	3.60	7.32	8.28		
δ(^{13}C) / ppm	78.3	72.2	78.5	63.4	63.4	119.5	128.6	164.0	145.1

Supplementary Figure 3

HSQC signals Peak C

δ(^{1}H) / ppm	3.75	3.74	3.82	3.4	3.8	3.78	3.58	3.68	3.66
δ(^{13}C) / ppm	62.9	63.1	63.1	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.21	3.76	5.8	5.42	7.28	8.23
δ(^{13}C) / ppm	76	79.1	99.1	102.4	119.4	128.7

HSQC signals pNP-maltose

δ(^{1}H) / ppm	3.76	3.74	3.83	3.41	3.81	3.8	3.57	3.68	3.67
δ(^{13}C) / ppm	62.9	63.2	63.2	72.0	73.5	74.1	74.4	75.4	75.5

δ(^{1}H) / ppm	4.22	3.77	5.82	5.43	7.31	8.27
δ(^{13}C) / ppm	76.1	79	99.1	102.4	119.5	128.7

Supplementary Figure 4